

Article

# Effects of Modeling Instruction Professional Development on Biology Teachers' Scientific Reasoning Skills

**Andria N. Stammen** <sup>1,\*</sup> , **Kathy L. Malone** <sup>2</sup>  and **Karen E. Irving** <sup>1</sup> <sup>1</sup> Department of Teaching and Learning, College of Education and Human Ecology, The Ohio State University, Columbus, OH 43210, USA; irving.8@osu.edu<sup>2</sup> Graduate School of Education, Nazarbayev University, Astana 01000, Kazakhstan; klmalone60@gmail.com

\* Correspondence: stammen.52@osu.edu

Received: 1 July 2018; Accepted: 6 August 2018; Published: 8 August 2018



**Abstract:** International assessments have revealed that students in numerous nations lack scientific reasoning skills. Science teachers who support students' scientific skill development through the use of authentic practices provide students with tools needed for success in future science courses. Teachers training focused on pedagogy that supports student scientific reasoning development is particularly important as some studies have also suggested that pre-service teachers have a tendency to display a lack of scientific reasoning skills. Additionally, few studies exist that assess teachers' scientific reasoning skills, including the effectiveness of professional development to strengthen teacher scientific reasoning abilities over time. To help fill this gap, this study examines the effects of a Modeling Instruction in a biology workshop on teachers' scientific reasoning skills. In addition to teacher interviews, focus groups, and writing samples, data from Lawson's Classroom Test of Scientific Reasoning (LCSR) were collected from teachers before and after the workshop. The results suggest that the three-week Modeling Instruction in the biology workshop contributed to gains in in-service teachers' scientific reasoning, and thus provides evidence that the teachers in this study are more prepared to help develop similar skills with their own students as they engage in the Modeling Instruction curriculum.

**Keywords:** scientific reasoning; in-service teachers; Modeling Instruction; LCSR; biology teachers

---

## 1. Introduction

International testing over the last decades has revealed that students in numerous nations do not have an adequate grasp of conceptual scientific knowledge nor the scientific reasoning skills that allow them to construct arguments based on scientific evidence [1,2]. The Organisation for Economic Co-operation and Development (OECD) [2] recommends the use of authentic practices, which include model-based pedagogy, to improve student ability. Students worldwide could benefit from authentic instruction that focuses on improving their scientific reasoning abilities. The development of student scientific reasoning skills has been correlated with increases in conceptual scientific knowledge in physics [3–5] and biology [6], as well as general science and mathematics [7]. Therefore, it is important that science teachers focus on authentic practices that allow students to develop their scientific reasoning skills in order to succeed in future science courses. Unfortunately, training of pre-service and in-service teachers may be lagging behind this need.

A few studies have assessed the scientific reasoning skills of pre-service science teachers, however, the results are not encouraging. Hilmert-Rüppell et al. [8] found that pre-service science teachers in Germany lacked a number of scientific reasoning skills, such as control of variables. In addition, it has

been found that pre-service teachers have a low understanding of biological models, which are essential when using the authentic practice of modeling [9], as well as the pedagogical content knowledge (PCK) to engage students [10]. Thus, the training of pre-service biology teachers in authentic practices, such as modeling, that enhance scientific reasoning has not been shown to be effective.

In addition, it has been shown that in-service teachers' knowledge of how to adapt their subject matter knowledge to engage diverse learners in scientific endeavors (i.e., PCK) affects the level of authentic practice utilized in classrooms [11]. If the level of authentic practice, such as modeling and multiple representation utilization, are not emphasized, Tsui and Treagust, [12] found that reasoning in genetics can be negatively impacted. In-service teachers have also shown that their understanding of biological models is limited [13,14]. Additionally, in-service teachers realize that their scientific reasoning skills are lacking as a majority of teachers self-select that they need to improve in their ability to engage in scientific inquiry [15]. Therefore, it is surprising that few studies exist that assess teacher scientific reasoning skills, as well as the effectiveness of professional development to strengthen teacher scientific reasoning abilities.

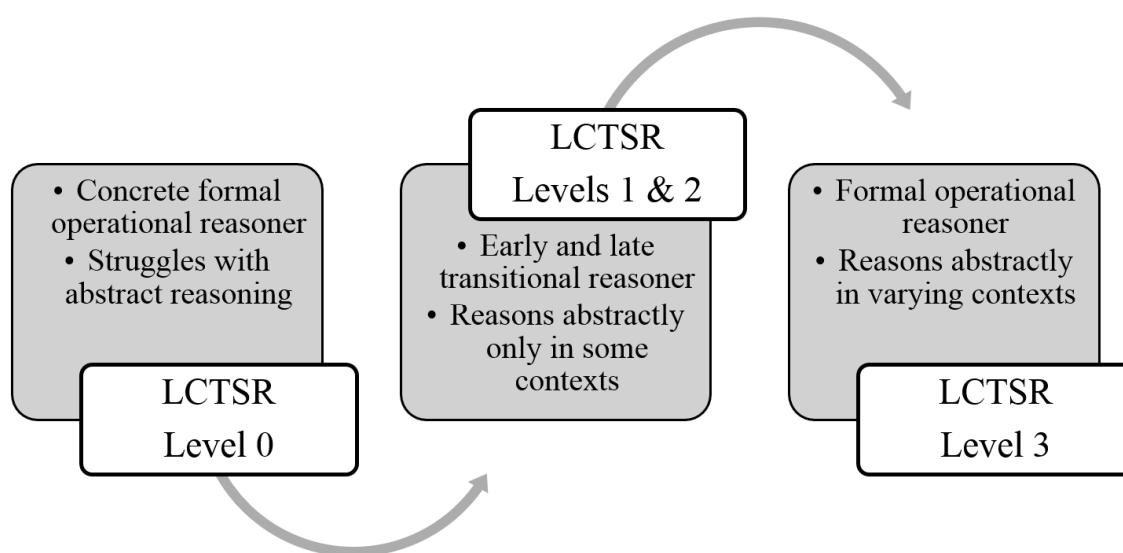
### 1.1. Scientific Reasoning

Reasoning is a process undertaken both formally in the academic domains, and informally, in everyday situations for making sense of phenomena, events, and processes. Reasoning has been identified as a skill necessary for effective engagement and learning of the theoretical and procedural knowledge associated with science disciplines [16,17]. In scientific efforts, reasoning is regarded as a key skill to ensure effective learning and engagement in all disciplines. Previous research has defined scientific reasoning in various ways. Some researchers define scientific reasoning as thinking about a topic and with scientific knowledge [18,19]. While this definition is a good starting point, it can be expanded to include subskills associated with scientific reasoning (i.e., hypothetical-deductive reasoning, control of variables, hypothesis generation, and evidence-based conclusion-making) [20].

Scientific reasoning ability has been regarded as crucial to science education. Recent PISA results have shown that only 7.7% of all students worldwide can evaluate methods to explore questions scientifically and identify limitations in experiments [2]. As a result, science education and research institutions, such as the Institute of Educational Assessors (IEA), the United Nations Educational, Scientific, and Cultural Organization (UNESCO), the American Association for the Advancement of Science (AAAS), and the National Academy of Sciences (NAS), have advocated, developed, and promoted national and international content standards and curriculum with scientific reasoning as a universal theme [21–23]. Furthermore, scientific reasoning has been identified as necessary for the "practice" of science in a number of countries [24,25]. These organizations recommend that the learning of scientific reasoning and scientific procedural/conceptual knowledge should be integrated. This recommendation seems appropriate as a number of studies have shown a connection between scientific reasoning ability and content knowledge gains. Specifically, scientific reasoning has been found to be a predictor of student performance in multiple contexts, especially when inquiry oriented engagement methods are utilized [3,6,26–29]. This curricular emphasis on the development of students' reasoning abilities implies that teachers may also need to be competent scientific reasoners in order to help their students develop their reasoning capacities. However, few studies have been conducted to determine secondary teachers' scientific reasoning abilities.

Many instruments have been employed to measure students' scientific reasoning, however, the majority tend to target a broad context of scientific literacy such as the nature of science and ethics in sociocultural settings [30]. Examples of such instruments include the Scientific Reasoning Test Version 9 (SR-90 [31] and Test of Scientific Literacy Skills (TOSLS) [32]. The Lawson's Classroom Test of Scientific Reasoning (LCTSR) [33], on the other hand, draws on the practices of scientific inquiry, such as hypothesis formulation and experimentation, and claims to measure scientific reasoning [30,33]. This LCTSR instrument has been utilized for nearly three decades and represents a task-based assessment that has been used to collect quantitative data on students' scientific reasoning

ability and skill development from middle school to college level [30,34]. According to Lawson ([33], p. 12), the purpose of the LCTSR is to “(1) measure concrete- and formal-operational reasoning; (2) be capable of administration to classes of secondary school and college age students in a relatively short period of time; (3) be easily scored; (4) use a format involving physical materials and require as little reading and writing as possible; and (5) include a large enough number and variety of problems to assure a high degree of reliability.” In 2000, Lawson released a modified version of the LCTSR that contained 24 multiple choice questions displaying the same two-tier design as the original version of the LCTSR. Generally, the first tier targets a “what” explanation and the second tier elicits a “why” explanation. The paired questions assess specific subskills of scientific reasoning. The LCTSR targeted the following scientific reasoning subskills: (i) matter and volume conservation, (ii) proportional reasoning, (iii) control of variables, (iv) probability reasoning, (v) correlation reasoning, and (vi) hypothetical-deductive reasoning [30,34]. Additionally, Lawson [33,34] incorporated the concrete operational, transitional, and formal operational levels of reasoning when developing the LCTSR. As shown in Figure 1, the LCTSR questions can be used to identify learners as Level 0 (formal operational reasoner), Levels 1 and 2 (transitional reasoner), or Level 3 (formal operational reasoner) [35]. The scoring levels are based on a two-tier scoring method and include the following ranges: 0–4 (concrete formal operational reasoners); 5–7 (early transitional reasoners); 8–10 (late transitional reasoners) and; 11–13 (formal operational reasoners) [7].



**Figure 1.** Comparison of the Lawson's Classroom Test of Scientific Reasoning's (LCTSR's) four levels of reasoning and the Piagetian cognitive development three levels of formal reasoning. Using a two-tier scoring method, LCTSR items can be analyzed to identify learners as Level 0: formal operational reasoner, Levels 1 and 2: transitional reasoner, or Level 3: formal operational reasoner whereby the following score ranges are utilized: 0–4 (concrete formal operational reasoners); 5–7 (early transitional reasoners); 8–10 (late transitional reasoners) and; 11–13 (formal operational reasoners) [7,35].

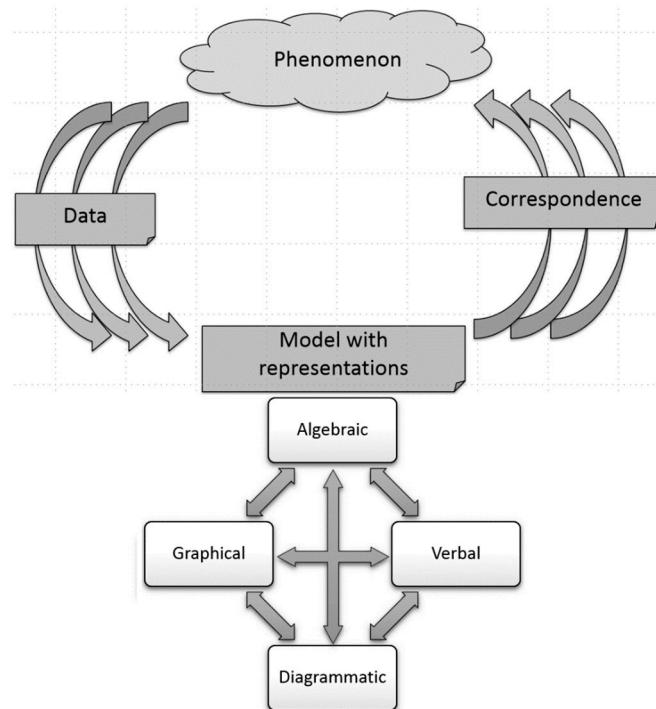
In this study, the instrument chosen to assess teachers' scientific reasoning was Lawson's Classroom Test of Scientific Reasoning (LCTSR) because the instrument seeks to measure scientific reasoning skills grounded in scientific inquiry and has been used with adult as well as secondary school age populations, thus allowing for comparisons between various populations of learners. In addition, the skills assessed by the LCTSR are aligned with ontological tenets of Modeling Instruction (MI).

### 1.2. Modeling Instruction (MI)

Studies have shown that students who are taught by using inquiry-based methods have demonstrated increases in their scientific reasoning skills [36–38]. However, the teaching methods

practiced in the world are failing to yield the desired results on PISA given that only 7.7% of students displayed advanced scientific literacy abilities [2]. Likewise, more than 90% of secondary students in the US unsuccessfully displayed formal operational reasoning skills as measured by the LCTSR [39]. To promote and improve students' scientific literacy, OECD [2] recommends the use of authentic practices in science education. One authentic practice that makes use of inquiry-oriented activities is Modeling Instruction (MI).

Modeling Instruction is a pedagogical technique that was developed in the mid-1990s for the teaching of physics [40] and since has expanded to other science disciplines including biology. Modeling Instruction is constructivist in nature and makes use of a learning cycle during instruction [41]. This instructional technique was designed to simulate authentic practices in science because scientists routinely construct models and use them within the context of modeling cycles when exploring and creating knowledge [42]. Thus, MI uses a learning cycle approach similar to the three phased cycle developed by Karplus [43] based on authentic science practices and Piagetian development theories. MI, however, collapses Karplus's three phases of exploration, concept introduction, and concept application into two phases which are model development and model deployment. The model development phase allows for student exploration and initial concept introduction. The model deployment phase consists of further concept development as well as concept application. Within each of these phases, students undergo a modeling cycle (see Figure 2) that consists of constant generation or analysis of data to produce a model with multiple representations while checking the model's correspondence with real world phenomena. MI provides students with the opportunity to engage in the authentic scientific practices of designing and conducting experiments and collecting and analyzing data; and in turn, these inquiry-based experiences allow students to develop and refine models representing the scientific phenomena being studied. That is, student groups analyze the collected data to determine any patterns between the tested variables. Then, these patterns are used by the students to construct external representations of their developing model. Student representations of the model can take the form of a graph, an algebraic equation, a physical prototype, a diagrammatic or verbal explanation. Students engage in consensus meetings where the class arrives at a final model that takes into account the multiple representations developed by the groups through the use of dialogic argumentative reasoning. This forum allows for student expression of understanding including how each representation not only describes the phenomena, but can be predictive of the behavior of the phenomena as well. Thus, all of the representations encompass the scientific model that can be used to describe, explain, and/or make predictions about the natural phenomena [41,44]. Students learn that no single representation truly exemplifies a scientific phenomenon; and as such, it takes multiple representations to portray the entire depth of the phenomena [42,45]. As students use the model and its multiple representations to solve problems and make predictions in other contexts, they develop knowledge that all the representations work together to comprise the model, but that different representations work better under specific conditions [45,46]. Students' initial models allow them to predict what might occur in these new situations which is called a deployment. After each deployment of the model, students must consider if their predictions were correct; and if the predictions were not correct, then adjustments would need to be made to their model. Students develop the ability to fluidly switch between representations as required, which is a characteristic that has been identified as expert-like behavior in scientific professions [47]. Figure 2 illustrates the Modeling Cycle. As the figure suggests, students are constantly checking their model by using it to make predictions. If that model fails to make accurate predictions, then students must revise the model based on the new data and the cycle continues as shown.



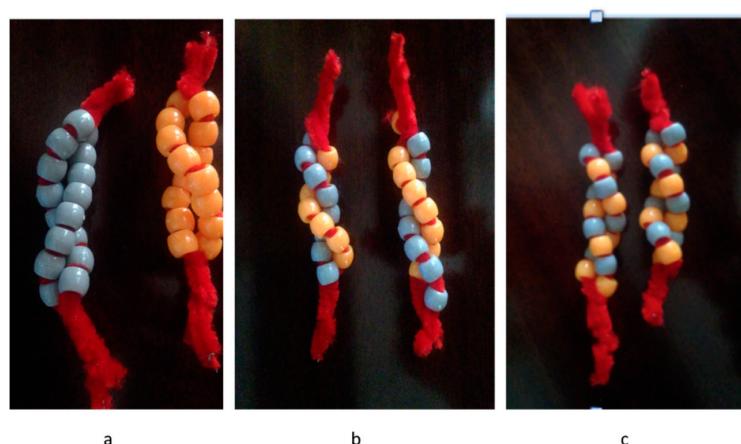
**Figure 2.** Modeling Cycle (adapted from [48]). This figure represents how Modeling Instruction (MI) explores scientific phenomenon through the use of data, models with representations in algebraic, graphical, diagrammatic, and/or verbal form, and correspondence.

For example, in the context of studying cell growth and reproduction, students develop the cycles of mitosis using the Modeling Cycle to determine that the number of chromosomes increases in some of the phases of mitosis but the final daughter cells consist of the same number of chromosomes as the parent cell. Students determine that the limitation of this model is that they do not understand how the chromosomal DNA replicates. Thus, students move towards the idea of developing a model for DNA replication to expand their current model of cell growth. When studying DNA replication (specifically, Meselson and Stahl's famous experiment), students construct a physical prototype of the replication process. In order to accomplish this learning, the students design physical models of "DNA" strands. One example would be using pop beads and chenille stems (see Figure 3). The materials in Figure 3 include a DNA strand with silver beads which represents a DNA strand created in the presence of N<sup>15</sup> isotopes. The orange beads represent N<sup>14</sup> isotopes that are used for the creation of new DNA by the students. Next, students determine possible methods to replicate the DNA during mitosis using the physical beads representing N<sup>14</sup> isotopes. Figure 4 shows three common replicated DNA strands that students develop. Students then compare their possible replication products to the results taken from the Meselson and Stahl density gradient centrifugation banding strips (see Figure 5 for an example). Students understand that the different densities of the DNA strands are due to variation in isotope composition and which cause the DNA strands to appear at different levels in a density gradient. Figure 4 is the density gradient centrifugation banding results of the first round of DNA replication showing that the real world results would display a mid-range density. Students discover that only the physical models from Figure 4 image b and image c match the data. The representations developed by each group are shared during a post lab consensus meeting where student groups discuss their findings and soon determine they cannot arrive at a consensus because there are two possible results that match Meselson and Stahl's lab data (see Figure 5). Therefore, students decide they need to try a second round of replication using the DNA strands shown in Figure 4 image b and image c. This final replication allows the class to reach the consensus that the physical model that best represents DNA replication is the process that led to the DNA strand in image b in which the

semi-conservative model of DNA replication is represented. Next, students represent their physical models with white-boarded diagrams and graphs (see Figure 6). At the final consensus meeting, the class arrives at a final model that takes into account the multiple representations developed by the groups through the use of dialogic argumentative reasoning. Thus, the students develop an understanding of the three DNA replication alternative models (i.e., conservative (Figure 4 image a), semi-conservative (Figure 4 image b), and dispersive (Figure 4 image c) without being told via direct instruction by the teacher. In the final step of this Modeling Cycle, students discuss the limitations of the physical DNA model and graphs and diagrammatic representations which lead students to the understanding that they do not know about the molecular structure of DNA (i.e., the nucleotide subunit comprised of a nitrogenous base, five-carbon sugar, and phosphate group).



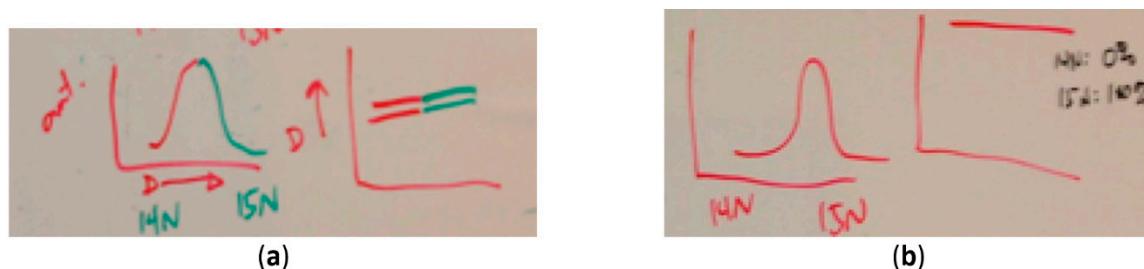
**Figure 3.** Example of physical model for DNA replication modeling using pop beads and chenille stems (silver beads = DNA containing N<sup>15</sup>; orange beads: DNA containing N<sup>14</sup>).



**Figure 4.** Possible DNA replication models developed by students (image (a) = conservative replication model; image (b) = semi-conservative replication model; image (c) = dispersive replication model).



**Figure 5.** Meselson and Stahl density gradient centrifugation banding strip displays the results after one round of DNA replication.



**Figure 6.** White-boarded graphical representations of students' physical models (i.e., beads and chenille stems). (a) semi-conservative replication model; (b) conservative replication model.

### Effectiveness of Modeling Instruction on Student Learning and Teacher Learning

The pedagogical practice of Modeling Instruction has been shown to have positive effects on secondary science students in terms of (i) increased use of expert-like problem solving and metacognitive skills [49], (ii) greater use of multiple representations [48,49], as well as (iii) improvements in conceptual performance [40,41,48–52], and (iv) scientific reasoning skills across biology, chemistry, and physics [53]. Malone, Schuchardt, and Sabree [48] discovered that biology students who were taught using a natural selection Modeling Instruction (MI) unit showed significantly higher gains in conceptual knowledge over that of a comparison group. In addition, the MI biology students also used a greater number of multiple representations when describing evolution. Comparable results were found with physics students who were taught using MI [49]. While solving complex physics problems during talk-aloud interviews, MI trained students were able to move fluidly between representations to check their work, whereas, the comparison group used mostly one representation to solve the problems and were more likely unable to check their solutions.

It becomes essential that we determine how to best train our science teachers given that students are often not being taught how to reason scientifically according to large scale testing like PISA. Given the lack of greater scientific literacy in our youth, it is possible that students are not being taught by using reformed science teaching methods that utilize inquiry-based methods and authentic practices in science. Studies of science teaching practice show that teachers continue to teach the way they were taught using direct instruction and rewarding memorizing and repeat performance [54]. Studies have also shown that pre-service teachers taught using inquiry based reform methods display increases in the use of such methods in their in-service instruction [55] and gains in their own scientific reasoning skills [56,57]. Adamson et al. [55] enrolled pre-service biology teachers into reformed college biology classes using inquiry methods and tracked their subsequent use of reformed methods as in-service teachers. The study revealed that when these teacher candidates became in-service teachers, they demonstrated an increased used of reformed methods and their students also had higher gains in scientific reasoning than comparison students. However, little research exists concerning the effects of an authentic practice like Modeling Instruction on either pre-service or in-service teachers.

#### 1.3. Quality Teacher Workshop Characteristics

Teaching in 21st century schools requires many skills. Teachers face challenges meeting the needs of diverse learners, collaborating with school administrators and colleagues, adapting to changes in disciplinary standards and expectations, withstanding scrutiny from the public and parents, staying abreast of relentless advances in technology, and understanding and implementing the constant changes required as a consequence of decisions in the legal groundings of schools. Once teachers join the work force, their need to continually refresh their skills is an ongoing challenge. Professional development is a strategy employed by many teachers, professional organizations, districts, and universities to enhance and strengthen teacher practice [58].

Characteristics of effective professional development have been researched by many. Loucks-Horsley, Stiles, Mundry, Love and Hewsen [59] identified the importance of clearly defined

goals that are focused on student learning. Professional development should consider teachers' knowledge, beliefs, and attitudes, and is most effective when closely tied to the needs of the professionals enrolled in the programs [59,60]. A variety of researchers emphasize the need to base professional development experiences on specific content and classroom practices [61,62]. PD providers are advised to plan a coherent outline of practices to develop both content and pedagogical knowledge concurrently about both what to teach and how students learn that particular content [63]. PD leaders should have appropriate content expertise and classroom experience to lead and collaborate with their professional students [59,63]. Active learning where teacher participants alternate roles between students and teachers with concrete tasks involving examining student work and opportunity to try out lesson plans and ideas have shown positive results in PD experiences [62,64]. Sustained and intensive experiences with follow-up and learning community supports have also been identified in research studies as key characteristics in effective PD (e.g., [58,61,62,65–67]).

Professional development workshops in Modeling Instruction for teachers have been conducted for many years [50]. These workshops were designed based on professional development research. While the effects of MI on students have been publicized to a great extent the effects of MI workshops on teachers has not been assessed to the same depth. Jackson et al. [50] published qualitative survey results showing that teachers self-reported that the MI workshops highly influenced the way they taught. However, there have been no published reports to date concerning how these workshops might affect science teachers' own scientific reasoning skills, problem solving skills or content knowledge. This study fills this gap by studying the effects of MI workshops on biology teachers' scientific reasoning skills.

## 2. Materials and Methods

### 2.1. Workshop Context

The workshop was designed with the elements of effective PD taken into account in order to support teachers as they designed experiments, gathered data, and analyzed that data to construct biology models. The teachers participated in the workshop in both student mode and teacher mode. In student mode, they worked through the modeling activity as their own students would. Also while in student mode, the teacher leaders took on the role of the classroom guide and helped the teachers to ask questions of others and interact as they constructed the scientific models in question. The teacher leaders were well versed in MI having taught with the pedagogy at the secondary level for three to ten years, as well as being part of the MI in biology development team for two years prior to the workshop. In teacher mode, the participants discussed how the enactment of these units would occur in their classrooms so that they could increase their self-efficacy concerning the implementation of the MI in biology units with their own students. For example, the teachers would practice developing questioning strategies to use with their own students. Engaging in these activities may have aided teachers in further developing and refining their scientific reasoning skills. Days 1–3 of the workshop focused on the unit of What is Life? Days 3–7 dealt with the unit of Population Interactions, and days 8–14 centered on Evolution. Teachers completed the pre-LCTSR on day 1 and post-LCTSR on day 14. Refer to Figure 7 for an overview of each workshop day's activities including the intended student learning objectives and the teacher modeler scientific reasoning goals.

<b>Day 1</b>
<ul style="list-style-type: none"> <li><b>Activities:</b> Pre-LCTSR; Characteristics of Life WB &amp; C</li> <li><b>Student Learning Objective:</b> Analyze what characterizes an organism as living.</li> <li><b>Modeler Scientific Reasoning Goal:</b> Engage in argumentative reasoning, correlational reasoning, and hypothetical-deductive reasoning.</li> </ul>
<b>Day 2</b>
<ul style="list-style-type: none"> <li><b>Activities:</b> Are Viruses Living or Nonliving Debate</li> <li><b>Student Learning Objective:</b> Analyze what characterizes an organism as living</li> <li><b>Modeler Scientific Reasoning Goal:</b> Engage in argumentative reasoning, correlational reasoning, and hypothetical-deductive reasoning.</li> </ul>
<b>Day 3</b>
<ul style="list-style-type: none"> <li><b>Activities:</b> Organism Classification Card Sort</li> <li><b>Student Learning Objective:</b> Based on morphological evidence, analyze the classification of organisms according to their evolutionary relationships.</li> <li><b>Modeler Scientific Reasoning Goal:</b> With minimal data, develop classification schemes using argumentative reasoning, correlational reasoning, and hypothetical-deductive reasoning.</li> </ul>
<b>Day 4</b>
<ul style="list-style-type: none"> <li><b>Activities:</b> Bacteria Lab; Paramecium Population Graphing Lab</li> <li><b>Student Learning Objective:</b> Design or simulate a population growth model which allows for the manipulation of environmental conditions.</li> <li><b>Modeler Scientific Reasoning Goal:</b> Analyze data using proportional reasoning, correlational reasoning, probabilistic reasoning, conservation of matter reasoning, hypothetical-deductive reasoning, and control of variable reasoning.</li> </ul>
<b>Day 5</b>
<ul style="list-style-type: none"> <li><b>Activities:</b> Misconception Article Reflection</li> <li><b>Modeler Scientific Reasoning Goal:</b> Analyze article using correlational reasoning and hypothetical-deductive reasoning.</li> </ul>
<b>Day 6</b>
<ul style="list-style-type: none"> <li><b>Activities:</b> Mathematical Modeling and Population Growth Lab WB &amp; C</li> <li><b>Student Learning Objective:</b> Apply mathematical models to explain homeostasis and evaluate population growth in ecosystems.</li> <li><b>Modeler Scientific Reasoning Goal:</b> Engage in argumentative reasoning; and analyze data using proportional reasoning, correlational reasoning, probabilistic reasoning, hypothetical-deductive reasoning, conservation of matter reasoning, and control of variable reasoning.</li> </ul>
<b>Day 7</b>
<ul style="list-style-type: none"> <li><b>Activities:</b> Population Growth Simulations &amp; WB; What is Engineering? Discussion</li> <li><b>Student Learning Objective:</b> Predict population interactions within an ecosystem based on the developed model of population growth.</li> <li><b>Modeler Learning Objective:</b> Analyze data using proportional reasoning, correlational reasoning, probabilistic reasoning, hypothetical-deductive reasoning, conservation of matter reasoning, and control of variable reasoning.</li> </ul>
<b>Day 8</b>
<ul style="list-style-type: none"> <li><b>Activities:</b> Lazy Lizards Natural Selection Lab</li> <li><b>Student Learning Objective:</b> Use data to determine relationships between a change in a population and its environment.</li> <li><b>Modeler Scientific Reasoning Goal:</b> Analyze data using proportional reasoning, correlational reasoning, probabilistic reasoning, hypothetical-deductive reasoning, and control of variable reasoning.</li> </ul>
<b>Day 9</b>
<ul style="list-style-type: none"> <li><b>Activities:</b> Lazy Lizards Natural Selection Lab WB &amp; C</li> <li><b>Student Learning Objective:</b> Summarize and analyze the survival and reproductive success of organisms in terms of behavioral, structural, and reproductive adaptations.</li> <li><b>Modeler Scientific Reasoning Goal:</b> Engage in argumentative reasoning; and analyze data using proportional reasoning, correlational reasoning, probabilistic reasoning, and hypothetical-deductive reasoning.</li> </ul>
<b>Day 10</b>
<ul style="list-style-type: none"> <li><b>Activities:</b> Field Trip to Conference on Evolution and Ecology Research</li> </ul>
<b>Day 11</b>
<ul style="list-style-type: none"> <li><b>Activities:</b> Lazy Lizards Natural Selection Lab C</li> <li><b>Student Learning Objective:</b> Develop a model that demonstrates how external biotic and abiotic conditions select for phenotypes that aid in survival.</li> <li><b>Modeler Scientific Reasoning Goal:</b> Engage in argumentative reasoning; and analyze data using proportional reasoning, correlational reasoning, probabilistic reasoning, and hypothetical-deductive reasoning.</li> </ul>
<b>Day 12</b>
<ul style="list-style-type: none"> <li><b>Activities:</b> Sexual Selection, Genetic Drift, and Speciation WB &amp; C</li> <li><b>Student Learning Objective:</b> Summarize and analyze the survival and reproductive success of organisms in terms of behavioral, structural, and reproductive adaptations.</li> <li><b>Modeler Scientific Reasoning Goal:</b> Engage in argumentative reasoning; and analyze data using proportional reasoning, correlational reasoning, probabilistic reasoning, and hypothetical-deductive reasoning.</li> </ul>
<b>Day 13</b>
<ul style="list-style-type: none"> <li><b>Activities:</b> Bioengineering Design Challenge Ideation</li> <li><b>Modeler Goal:</b> Develop an engineering design challenge in biology that target students' content knowledge and reasoning skill development.</li> <li><b>Modeler Scientific Reasoning Goal:</b> Engage in argumentative reasoning, correlational reasoning, and hypothetical-deductive reasoning.</li> </ul>
<b>Day 14</b>
<ul style="list-style-type: none"> <li><b>Activities:</b> Bioengineering Design Challenge Ideation; Closure; Post-LCTSR</li> <li><b>Modeler Scientific Reasoning Goal:</b> Engage in argumentative reasoning, correlational reasoning, and hypothetical-deductive reasoning.</li> </ul>

**Figure 7.** Instruction Chain of Modeling Workshop in biology (WB: white boarding; C: consensus building).

## 2.2. Workshop Participants (Modelers)

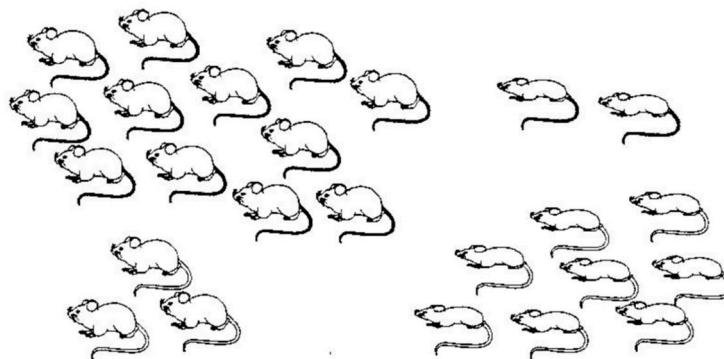
The workshop provided model-based biology curricula training to 32 secondary teachers (grades 6–12) employed at schools servicing students in both urban and suburban areas in the Midwest of the United States. Of these 32 teacher modelers, 30 completed the LCTSR before and after the workshop. Within this participant group, five were middle school educators, three were English language learner (ELL) co-educators, and 22 were high school educators.

## 2.3. Measures

### 2.3.1. LCTSR

The 24 item, selected response, two-tier designed Lawson's Classroom Test of Scientific Reasoning (LCTSR) was administered as both a pre- and post-assessment to the modelers who attended the three-week workshop [34] (see Figure 8 for a sample item). The LCTSR was constructed on a foundation of the Piagetian theoretical perspective of cognitive development. The theory of cognitive development is comprised of the following three formal reasoning levels: concrete operational, transitional, and formal operational [5,68]. A learner classified as a concrete operational reasoner displays logical skills in concrete tasks but finds difficulty when applying problem solving to abstract concepts [5]. A formal operational reasoner displays the ability to reason abstractly in varying contextual situations and thus displays the skills associated with "thinking like a scientist" ([5], p. 4). A transitional reasoner is an intermediate classification between the two fore-mentioned reasoners and demonstrates the ability to reason abstractly in some contexts [5]. The LCTSR questions can be used to identify learners as these various reasoner categories. The scoring levels are based on a two-tier scoring method (i.e., examining paired items collectively). The following ranges are used to identify reasoning levels using the LCTSR: 0–4 (concrete reasoners/formal operational reasoners); 5–7 (early transitional reasoners); 8–10 (late transitional reasoners) and; 11–13 (formal operational reasoners) [7]. Both the 1978 and 2000 versions of the LCTSR targeted specific scientific reasoning subskills. These subskills are as follows: (i) matter and volume conservation, (ii) proportional reasoning, (iii) control of variables, (iv) probability reasoning, (v) correlation reasoning, and (vi) hypothetical-deductive reasoning [33,34]. Proportional reasoning involves making sense of co-variant relationships that are expressed in different rates, units, and/or ratios. Correlation reasoning involves discerning the relationships between two or more variables. Hypothetical-deductive reasoning involves making sense of a general theory which includes the possible outcome and using deductive processes to apply that theory to make predictions/hypotheses to predict future events. Control of variables (COV) reasoning involves making sense of the relationships between independent and dependent variables. Reasoning within the context of matter and volume conservation are not discussed exhaustively in this paper since the Modeling Instruction workshop curriculum dealt primarily with the aforementioned reasoning skills since these skills were used consistently throughout all units of the curriculum.

Farmer Brown was observing the mice that live in his field. He discovered that all of them were either fat or thin. Also, all of them had either black tails or white tails. This made him wonder if there might be a link between the size of the mice and the color of their tails. So he captured all of the mice in one part of his field and observed them. Below are the mice that he captured.



*Do you think there is a link between the size of the mice and the color of their tails?*

- a. appears to be a link
- b. appears not to be a link
- c. cannot make a reasonable guess

*because*

- a. there are some of each kind of mouse.
- b. there may be a genetic link between mouse size and tail color.
- c. there were not enough mice captured.
- d. most of the fat mice have black tails while most of the thin mice have white tails.
- e. as the mice grew fatter, their tails became darker.

**Figure 8.** Example of a Lawson's Classroom Test of Scientific Reasoning (LCTSR) correlation reasoning two-tier item [33,34].

To better understand our teacher modelers, their performance on the LCTSR was investigated in two ways: single-tier item approach and two-tier item approach. The single-item analysis approach (i.e., treating all 24 items as independent) was used to determine the overall LCTSR performance scores and subskill average scores. Using the two-tier item analysis approach (i.e., treating paired items as a collective), the LCTSR scores were categorized based on reasoning classifications.

### 2.3.2. Teacher Focus Groups, Interviews, and Writing Samples

The teachers attending the Modeling Instruction workshop participated in focus group interviews on the last day of the summer workshop and individual interviews at the end of the school year following the implementation of the MI curriculum with their biology students for a full academic school year. All of the interviews were semi-structured meaning that not only was there a number of set of standard questions used for both interviews, but also the interviewees' comments guided the direction of the interviews including the types of follow-up questions asked. Each of the six focus groups included approximately five teacher modelers. The interviews lasted approximately 60 min and included 17 questions. Eleven modelers volunteered to participate in the end of the school year individual interviews which lasted between 30 and 60 min and included twenty six questions. All interviews were audio taped and transcribed. Out of the 26 questions asked during both interviews, the responses to three questions that were the same for both interviews are reported in this paper and are as follows:

1. After the workshop, do you have new ways to help your students experience biology topics in your classroom?
2. What do you think your students will learn about biology from incorporating models into your classroom?
3. What do you think your students will learn about biology from incorporating modeling instruction into your classroom?

These selected interview questions do not specifically ask information about scientific reasoning. During the MI workshops scientific reasoning skills were taught implicitly via the classroom activities when teachers were in student mode. The questions were used to determine if the teachers made the explicit connections between the MI methodology and the targeting of scientific reasoning skills. The interviews were audio recorded and transcribed. The transcriptions were checked for accuracy by the researchers. The interview transcripts were read and coded using a coding scheme developed for this project that focused on the scientific reasoning sub-skills. The sub-skills included control of variables, proportional reasoning, probability reasoning, correlation reasoning, and hypothetical-deductive reasoning. Before coding the answers to the above questions, responses were segmented into coherent thought segments. The subset of the interviews was read and a code book was constructed with example statements for each code. The open coding included segmenting the answers to the above questions into coherent thought fragments. These segments were then coded using the coding scheme for references to scientific reasoning skills (see Table 1 for an example of the codebook).

**Table 1.** Examples from the codebook.

Main Code	Examples Coded from Interviews
Control of Variables	FG1: start with research question and generate data FG4: how to gather data
Proportional Reasoning	FG1: look for patterns that represent the data
Hypothetical-deductive reasoning	FG2: a process of thinking FG3: [conclusions are] evidence based; evidence and reasoning

At the end of the summer workshop (post sample) and the academic school year (delayed post sample), teachers were asked to answer a single question pertaining to Modeling Instruction. This writing sample focused on one question: what does modeling in science mean? These writing samples were segmented and coded using the same codebook described above and in Table 1.

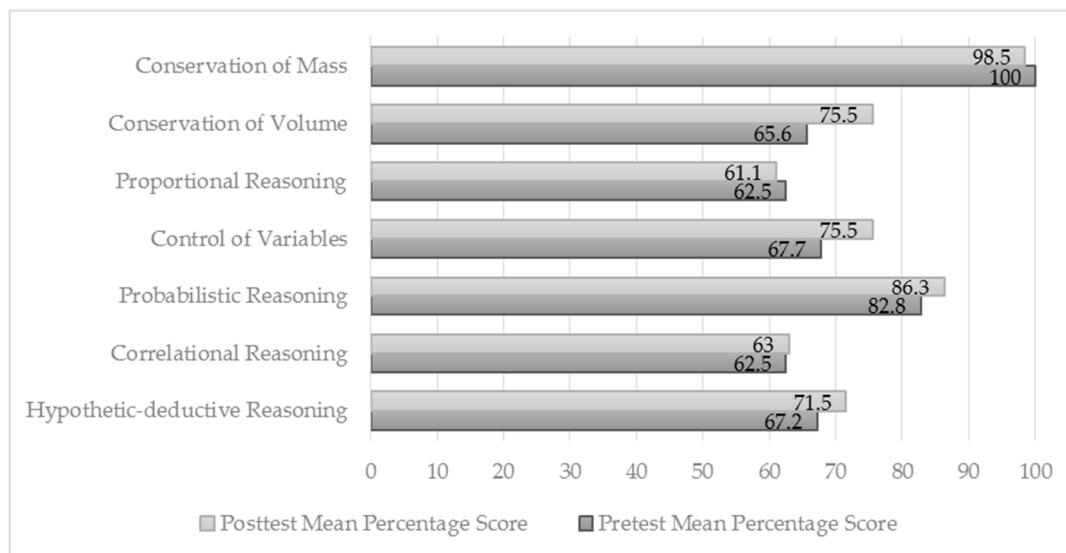
### 3. Results

#### 3.1. Single-Tier Item Analysis

The single-tier item analysis approach was used to determine the overall LCTSR and subskill average scores. To determine the overall average score, each of the 24 questions on the LCTSR was treated as a stand-alone question. On average, the teacher modelers correctly answered 17 of the 24 questions on the pretest, which equates to a mean score of 71.4% ( $N = 30$ ,  $SD = 3.97$ ). Figure 5 shows the sample's mean percentage score for specific reasoning subskills assessed by the LCTSR. Modelers demonstrated competency within the subskills of conservation of matter (100%) and probabilistic reasoning (82.8%) but seemed to score lower on items related to correlational (62.5%) and proportional reasoning (62.5%) subskills.

On average, the teachers correctly answered 18 of the 24 questions on the post-test, which equates to a mean score of 75.0% ( $N = 30$ ,  $SD = 4.41$ ). Given that raw scores for assessments are seldom linear nor of equal difficulty, Rasch person measures were utilized when comparing mean performance on the LCTSR. A paired-samples t-test was conducted to compare pre-test performance and post-test performance. There was a significant difference in the scores for the pre-test ( $M = 1.24$ ,  $SD = 1.12$ )

and post-test ( $M = 1.64$ ,  $SD = 1.45$ ) suggesting that three-week Modeling Instruction in the biology workshop contributed to gains in in-service teachers' scientific reasoning ( $t(29) = -2.74$ ,  $p < 0.011$ ). Figure 9 shows the averages for specific reasoning subskills measured by the LCTSR during the post-test. Again, modelers performed better within the subskills of conservation of mass (98.5%) and probabilistic reasoning (86.3%) but seemed to perform lower on the correlational (63.0%) and proportional reasoning (61.1%) subskills. All subskills mean scores increased during the post-test except for conversation of mass and proportional reasoning. Although conversation of mass was not a central focus of the MI workshop, the proportional reasoning subskill was a scientific reasoning goal during several days of the workshop; and therefore, these results suggest that the teacher modelers need more practice making sense of co-variant relationships that are expressed in different rates, units, and/or ratios.



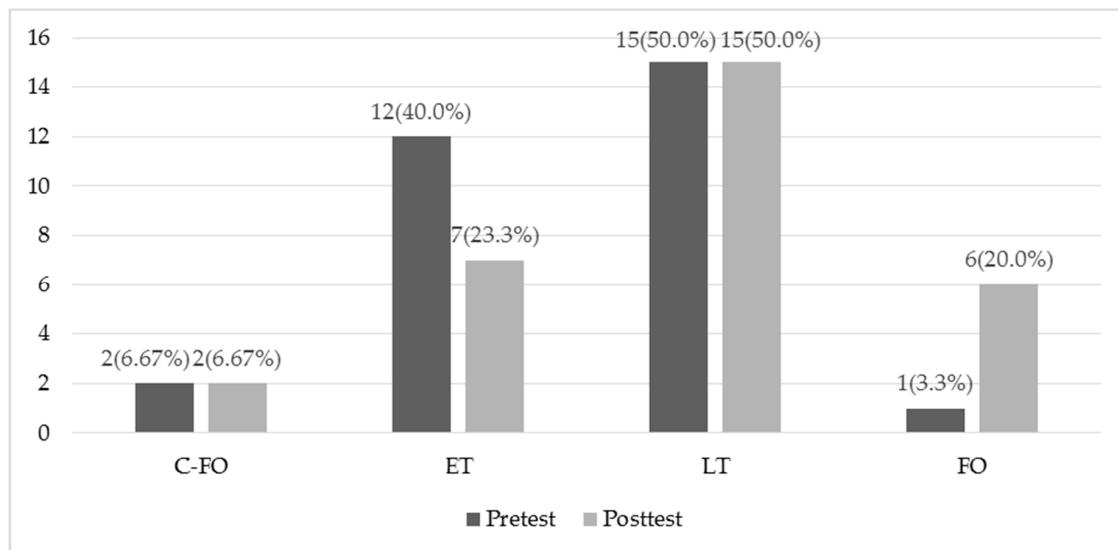
**Figure 9.** Pretest and post-test mean percentage scores on Lawson's Classroom Test of Scientific Reasoning (LCTSR) subskills from the single-tier item analysis showing increased post-test scores in all subskills except for conversation of mass and proportional reasoning.

### 3.2. Two-Tier Item Analysis

The two-tier item analysis approach was used to determine the distributions of the formal reasoning categories. Using individual teacher scores on the LCTSR, modelers were classified into one of the four formal reasoning categories. Unlike the single-tier item analysis approach which considered the correctness of each individual item, the two-tier item analysis approach requires a teacher to correctly answer both items within a scenario correctly to receive credit. Using this analysis technique, teacher modelers correctly answered on average 7 of the 12 question-pairs on the pretest, which equates to a mean score of 61.3% ( $N = 30$ ,  $SD = 2.08$ ). The following scoring levels were used to classify the teachers based on scientific reasoning ability: 0–4 (C-FO: concrete reasoners/formal operational reasoners); 5–7 (ET: early transitional reasoners); 8–10 (LT: late transitional reasoners) and; 11–13 (FO: formal operational reasoners) [7]. Figure 6 shows the distribution of the formal reasoning categories assessed by the LCTSR. Most teachers were classified as late transitional reasoners ( $N = 15$ ), followed by the category of early transitional researchers ( $N = 12$ ). Only one teacher was identified as a formal operational reasoner.

On the post-test, teachers correctly answered, on average, 9 of the 12 question-pairs, which equates to a mean score of 75.0% ( $N = 30$ ,  $SD = 2.69$ ). Figure 10 shows the population distribution of the formal reasoning categories assessed by the LCTSR. The majority of teachers were again classified as late transitional reasoners ( $N = 15$ ), followed by the category of early transitional researchers ( $N = 7$ ).

Six modelers were identified as a formal operational reasoner based on post-test data. This shift towards late transitional reasoners and formal operational reasoners suggest that the teacher modelers are more capable of reasoning abstractly in varying contexts following the Modeling Instruction in the biology workshop.



**Figure 10.** Pretest and post-test formal reasoning level distribution (C-FO: concrete formal operational reasoners with a score range from 0–4; ET: early transitional reasoners with a score range from 5–7; LT: late transitional reasoners with a score range from 8–10; and FO: formal operational reasoners with a score range from 11–13) [7].

### 3.3. Focus Groups, Interviews, and Writing Sample

Of the six focus groups, five groups discussed scientific reasoning subskills without prompting when asked the three open ended questions. In addition, 6 out of 11 modelers spoke about scientific reasoning sub-skills during the individual interviews at the end of the school year. The transcription comments were coded in terms of the sub-skill that was discussed. The majority of the statements focused on the scientific reasoning skills learned through Modeling Instruction. Most of the scientific reasoning statements focused on control of variables and hypothetical-deductive reasoning, 47% and 33% of the coded statements, respectively. For example, when speaking about control of variables one teacher said that MI focused on “how research is done and how scientists test ideas.” An example of hypothetical-deductive reasoning was contained in the following comment about MI “start with...a research question, you generate some data, you look for the patterns that represent the data, and then you come up with a model.” This statement included the sub-skill of proportional reasoning when it focused on determining patterns in the data. Another Modeler stated, “student centered lessons that engage the student to learn and utilize active thinking skills”, which suggests that MI promotes students to participate in scientific reasoning skills.

Based on the delayed post interviews, the implementation of MI pedagogy seemed to support the development of scientific reasoning skills in teachers and students alike. As one teacher said during the delayed post interview, Modeling Instruction is about “placing each learner in the role of a scientist and letting them explore and learn about their world by experiencing real science situations.” It is this focus of Modeling Instruction pedagogy which supports scientific reasoning shifts in participants. Another in-service teacher expressed that Modeling Instruction gets participants to “think like scientists by engaging them in observation, analysis, drawing conclusions, reaching consensus, then applying the consensus model to new or different situations. If the model holds up great! If not, then a second iteration is needed.”

The analysis of the writing sample showed comparable results in both the post and delayed post writing sample with the majority of the statements focused on MI's use of control of variables (38% vs. 28%, respectively) and hypothetical-deductive reasoning (38% vs. 28%, respectively). However, in the delayed post writing sample, teachers wrote more often of the use of proportional, probability, and correlational reasoning (10% in the delayed post vs. 8% in the post sample). Moreover, in both the interviews and writing samples, there was never a specific mention of scientific reasoning skills in general or specifically connected to any aspect of the Modeling Instruction workshop nor the use of MI units during the implementation with students. In fact, only one teacher mentioned that the students were learning skills, but even then, it was a generic mention that students were learning "a lot of good skills." In the writing sample, only two teachers mentioned skills but in a circumlocutory way such as "active thinking skills." Thus, teachers did not seem to explicitly connect the generation of scientific reasoning skills with MI pedagogy.

#### 4. Discussion

The single-tier item analysis indicated that scientific reasoning gains occurred between assessment administrations. The two-tier item analysis suggested shifts in teacher modelers' scientific reasoning towards late transitional and formal operational reasoning. This shift towards formal operational reasoning is indicative of the ability to reason abstractly in varying contextual situations as is displayed by scientists. These results suggest that the three-week Modeling Instruction in the biology workshop contributed to gains in in-service teachers' scientific reasoning making them more prepared to help develop similar skills within their own students while implementing Modeling Instruction. The interview results provide additional support to the LCTSR findings demonstrating shifts in scientific reasoning skills. The interview comments about modeling demonstrate that teachers were able to make the implicit connection within the Modeling Instruction units to the development of scientific reasoning skills. However, the teachers did not seem to make this connection explicitly since none mentioned scientific reasoning skills or reasoning in general in interviews or writing samples.

The greatest overall increase in sub-skills occurred within the context of control of variables (about a 9% increase pretest to post-test) and this sub-skill was the one most often mentioned by teachers in the interviews and writing samples. The second largest increase in subskill shown on the LCTSR was in hypothetical-deductive reasoning with about a 4% increase pretest to post-test. The Modelers engaged in hypothetical-deductive reasoning in 13 out of 14 of the MI workshop days, and this skill was also the second most often mentioned in the writing samples, interviews, and focus groups. However, the sub-skills of proportional and correlational reasoning were not impacted by the Modeling Instruction workshop even though Modelers engaged in these activities 7 and 13 days, respectively, out of the 14 days of the MI workshop. This is a discouraging finding and may highlight the need to incorporate explicit connections to proportional and correlational reasoning skills into future Modeling Instruction workshops through the use of additional deployment tasks that target each type of reasoning skill independently rather than concurrently.

#### 5. Conclusions

This study provides evidence that Modeling Instruction grounded in biology and authentic practice has the capacity to support the development of in-service teachers' scientific reasoning skills. After participating in the three-week workshop, teachers related how MI pedagogy enhances and promotes scientific reasoning subskills, especially as it relates to control of variable reasoning and hypothetical-deductive reasoning. In part, these results may be attributed to the development and implementation of the Modeling Instruction workshop which utilized the following effective practices associated with quality professional development: (i) concurrent exploration of content and pedagogy that is embedded in practice (i.e., modeling cycle in teacher-mode and student-mode); (ii) experienced workshop leaders with classroom and content experience; (iii) focused content (i.e., What is Life?, Population Interactions, Evolution units of study); (iv) sustained duration and with a professional

learning community (i.e., three-week workshop with follow-up meeting and interviews throughout the academic year); (v) planned with a coherent set of strategies (i.e., white boarding, consensus meetings, deployments) [58,61,62,65–67].

Although the modelers displayed gains in scientific reasoning skills, only 20% of the teachers were identified as formal operational reasoners at the end of the workshop. This result is not surprising however, as past research suggests that pre-service teachers have a tendency to display a lack of scientific reasoning skills [8]. When coupled, these results for pre-service and in-service teachers do not bode well for the education received during their college years and stress the need for in-service workshops that allow educators to experience authentic practices, as well as a need to explore shifts in pre-service teacher training. As has been found in the teaching and learning of nature of science (NOS) concepts and processes [69], scientific reasoning skills may not be easily apparent when taught implicitly. Thus, there is a need for teacher education to include opportunities for the explicit instruction of scientific reasoning skills which may include metacognitive reflection practices on scientific reasoning skill utilization after teachers engage in activities in student-mode.

While this study illuminates the scientific reasoning skills of a small group of life science teachers who engaged in this project, more studies are needed to extrapolate these findings to other science disciplines and other areas of the world. This study is limited by the small number of teachers who completed the assessments and participated in interviews and focus groups, as well as the fact that all the teachers were from one science discipline (biology).

**Author Contributions:** Conceptualization, K.L.M., K.E.I., A.N.S.; Methodology, K.L.M., A.N.S.; Validation, K.L.M., A.N.S.; Investigation, K.L.M., A.N.S.; Resources, K.L.M., K.E.I., A.N.S.; Writing-Original Draft Preparation, K.L.M., K.E.I., A.N.S.; Writing-Review & Editing, K.L.M., K.E.I., A.N.S.; Visualization, K.L.M., A.N.S.; Funding Acquisition, K.L.M.

**Funding:** This research was funded by a grant under the federally funded Math Science Partnership State Grants Program, administered by the Ohio Department of Education. Any opinions, findings, and conclusions or recommendations expressed in this material are those of the authors and do not necessarily reflect the views of the funding organizations.

**Conflicts of Interest:** The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript, and in the decision to publish the results.

## References

1. Kelly, D.; Xie, H.; Nord, C.W.; Jenkins, F.; Chan, J.Y.; Kastberg, D. *Performance of U.S. 15-Year-Old Students in Mathematics, Science, and Reading Literacy in an International Context: First Look at PISA 2012*; U.S. Department of Education, Washington's National Center for Education Statistics: Washington, DC, USA, 2013. Available online: <http://nces.ed.gov/pubs> (accessed on 8 October 2016).
2. Organisation for Economic Co-operation and Development (OECD). *What Are Tertiary Students Choosing to Study?* OECD Publishing: Paris, France, 2014; Available online: <http://www.oecd.org/education/skills-beyond-school/EDIF%202014--No19.pdf> (accessed on 24 June 2018).
3. Coletta, V.P.; Phillips, J.A.; Steinert, J.J. Why you should measure your students' reasoning ability. *Phys. Teach.* **2007**, *45*, 235–238. [[CrossRef](#)]
4. Ding, L. Verification of causal influences of reasoning skills and epistemology on physics conceptual learning. *Phys. Rev. Spec. Top. Phys. Educ. Res.* **2014**, *10*, 023101. [[CrossRef](#)]
5. Moore, J.C.; Rubbo, L.J. Scientific reasoning abilities of nonscience majors in physics-based courses. *Phys. Rev. Spec. Top. Phys. Educ. Res.* **2012**, *8*, 10106. [[CrossRef](#)]
6. Lawson, A.E.; Banks, D.L.; Logvin, M. Self-efficacy, reasoning ability and achievement in college biology. *J. Res. Sci. Teach.* **2007**, *44*, 706–724. [[CrossRef](#)]
7. Deming, J.C.; O'Donnell, J.R.; Malone, C.J. Scientific literacy: Resurrecting the phoenix with thinking skills. *Sci. Educ.* **2012**, *21*, 10–17.
8. Hilfert-Rüppell, D.; Looß, M.; Klingenberg, K.; Eghtessad, A.; Höner, K.; Müller, R.; Strahl, A.; Pietzner, V. Scientific reasoning of prospective science teachers in designing a biological experiment. *Lehrerbildung auf dem Prüfstand* **2013**, *6*, 135–154.

9. Reinisch, B.; Krüger, D. Preservice biology teachers' conceptions about the tentative nature of theories and models in biology. *Res. Sci. Educ.* **2018**, *48*, 71–103. [[CrossRef](#)]
10. Großschedl, J.; Harms, U.; Kleickmann, T.; Glowinski, I. Preservice biology teachers' professional knowledge: Structure and learning opportunities. *J. Sci. Teach. Educ.* **2015**, *26*, 291–318. [[CrossRef](#)]
11. Park, S.; Jang, J.Y.; Chen, Y.C.; Jung, J. Is pedagogical content knowledge (PCK) necessary for reformed science teaching? Evidence from an empirical study. *Res. Sci. Educ.* **2011**, *41*, 245–260. [[CrossRef](#)]
12. Tsui, C.; Treagust, D. Evaluating secondary students' scientific reasoning in genetics using a two-tier diagnostic instrument. *Int. J. Sci. Educ.* **2010**, *32*, 1073–1098. [[CrossRef](#)]
13. Krell, M.; Krüger, D. Testing models: A key aspect to promote teaching activities related to models and modelling in biology lessons? *J. Biol. Educ.* **2016**, *50*, 160–173. [[CrossRef](#)]
14. Ware, T.; Malone, K.L.; Irving, K.; Mollohan, K. Models and modeling: An evaluation of teacher knowledge. In Proceedings of the HICE 2017: The 15th Annual Hawaii International Conference on Education, Honolulu, HI, USA, 3–6 January 2017.
15. Zhang, M.; Parker, J.; Koehler, M.J.; Eberhardt, J. Understanding inservice science teachers' needs for professional development. *J. Sci. Teach. Educ.* **2015**, *26*, 471–496. [[CrossRef](#)]
16. Kuhn, D. Children and adults as intuitive scientists. *Psychol. Rev.* **1989**, *96*, 674–689. [[CrossRef](#)]
17. Lawson, A.E. The nature and development of scientific reasoning: A synthetic view. *Int. J. Sci. Math. Educ.* **2004**, *2*, 307–338. [[CrossRef](#)]
18. Hogan, K.; Fishkeller, J. Dialogue as data: Assessing students' scientific reasoning with interactive protocols. In *Assessing Science Understanding: A Human Constructivist View*; Mintzes, J.J., Wandersee, J.H., Novak, J.D., Eds.; Elsevier: Burlington, MA, USA, 2005; pp. 95–127, ISBN 0124983650 9780124983656.
19. Samarapungavan, A. Reasoning 2009. Available online: [www.education.com/references/article/reasoning](http://www.education.com/references/article/reasoning) (accessed on 14 November 2014).
20. Lawson, A.E.; Clark, B.; Cramer-Meldrum, E.; Falconer, K.A.; Sequist, J.M.; Kwon, Y.-J. Development of scientific reasoning in college biology: Do two levels of general hypothesis-testing still exist? *J. Res. Sci. Teach.* **2000**, *37*, 81–101. [[CrossRef](#)]
21. National Research Council. *A Framework for K-12 Science Education: Practices, Crosscutting Concepts, and Core Ideas*; National Academies Press: Washington, DC, USA, 2012.
22. Mullis, I.V.S.; Martin, M.O.; Loveless, T. *20 Years of TIMSS: International Trends in Mathematics and Science Achievement, Curriculum, and Instruction*; TIMSS & PIRLS International Study Center, Boston College: Chestnut Hill, MA, USA, 2016; ISBN 978-1-889938-40-0.
23. UNESCO. *Current Challenges in Basic Science Education*; UNESCO Education Sector: Paris, France, 2010.
24. NGSS Lead States. *Next Generation Science Standards: For States, by States*; The National Academies Press: Washington, DC, USA, 2013.
25. KMK [Sekretariat der Ständigen Konferenz der Kultusminister der Länder in der BRD]. *Bildungsstandards im Fach Biologie für den Mittleren Schulabschluss* [Biology education standards for the Mittlere Schulabschluss]; Wolters Kluwer: Neuwied, Germany, 2008.
26. Coletta, V.P.; Phillips, J.A. Interpreting FCI scores: Normalized gain, preinstruction scores and scientific reasoning ability. *Am. J. Phys.* **2005**, *73*, 1172–1182. [[CrossRef](#)]
27. Dauer, J.T.; Momsen, J.L.; Speth, E.B.; Makohon-Moore, S.C.; Long, T.M. Analyzing change in students' gene-to-evolution models in college-level introductory biology. *J. Res. Sci. Teach.* **2013**. [[CrossRef](#)]
28. Johnson, M.A.; Lawson, A.E. What are the relative effects of reasoning ability and prior knowledge on biology achievement in expository and inquiry classes? *J. Res. Sci. Teach.* **1998**, *35*, 89–103. [[CrossRef](#)]
29. Tfeily, O.; Dancy, M. Why do students struggle in freshman physics? In *2006 Physics Education Research Conference AIP Conference Proceedings*; McCullough, L., Hsu, L., Heron, P., Eds.; American Institute of Physics: College Park, MD, USA, 2007.
30. Ding, L.; Wei, X.; Mollohan, K. Does Higher Education Improve Student Scientific Reasoning Skills? *Int. J. Sci. Math. Educ.* **2016**, *14*, 619–634. [[CrossRef](#)]
31. Sundre, D.L.; Thelk, A.D. Advancing assessment of quantitative and scientific reasoning. *Numeracy* **2010**, *3*, 1–12. [[CrossRef](#)]
32. Gormally, C.; Brickman, P.; Lutz, M. Developing a test of scientific literacy skills (TOSLS): Measuring undergraduates' evaluation of scientific information and arguments. *CBE-Life Sci. Educ.* **2012**, *11*, 364–377. [[CrossRef](#)] [[PubMed](#)]

33. Lawson, A.E. The development and validation of a classroom test of formal reasoning. *J. Res. Sci. Teach.* **1978**, *15*, 11–24. [[CrossRef](#)]
34. Lawson, A.E. Classroom Test of Scientific Reasoning: Multiple Choice Version. *J. Res. Sci. Teach.* **2000**, *15*, 11–24. [[CrossRef](#)]
35. Lawson, A.E.; Alkhoury, S.; Benford, R.; Clark, B.R.; Falconer, K.A. What kinds of scientific concepts exist? Concept construction and intellectual development in college biology. *J. Res. Sci. Teach.* **2000**, *37*, 996–1018. [[CrossRef](#)]
36. Marshall, J.C.; Smart, J.B.; Alston, D.M. Inquiry-based instruction: A possible solution to improving student learning of both science concepts and scientific practices. *Int. J. Sci. Math. Educ.* **2017**, *15*, 777–796. [[CrossRef](#)]
37. Westbrook, S.L.; Rogers, L.N. Examining the development of scientific reasoning in ninth-grade physical science students. *J. Res. Sci. Teach.* **1994**, *31*, 65–76. [[CrossRef](#)]
38. Wilson, C.D.; Taylor, J.A.; Kowalski, S.M.; Carlson, J. The relative effects and equity of inquiry-based and commonplace science teaching on students' knowledge, reasoning, and argumentation. *J. Res. Sci. Teach.* **2010**, *47*, 276–301. [[CrossRef](#)]
39. O'Donnell, J.R. Creation of National Norms for Scientific Thinking Skills Using the Classroom Test of Scientific Reasoning. Master's Thesis, Winona State University, Winona, MN, USA, 2011.
40. Hestenes, D.; Wells, M.; Swackhamer, G. Force concept inventory. *Phys. Teach.* **1992**, *30*, 141–158. [[CrossRef](#)]
41. Wells, M.; Hestenes, D.; Swackhamer, G. A modeling method for high school physics instruction. *Am. J. Phys.* **1995**, *63*, 606–619. [[CrossRef](#)]
42. Hestenes, D. Modeling theory for math and science education. In *Modeling Students' Mathematical Modeling Competencies*; Springer: Boston, MA, USA, 2010; pp. 13–41, ISBN 978-1-4419-0561-1.
43. Karplus, R.; Butts, D.P. Science teaching and the development of reasoning. *J. Res. Sci. Teach.* **1977**, *14*, 169–175. [[CrossRef](#)]
44. Buckley, B.C.; Gobert, J.D.; Kindfield, A.C.H.; Horwitz, P.; Tinker, R.F.; Gerlits, B.; Wilensky, U.; Dede, C.; Willett, J. Model-based teaching and learning with BioLogica™: What do they learn? How do they learn? How do we know? *J. Sci. Educ. Technol.* **2004**, *13*, 23–41. [[CrossRef](#)]
45. Giere, R.N. How models are used to represent reality. *Philos. Sci.* **2004**, *71*, 742–752. [[CrossRef](#)]
46. Godfrey-Smith, P. The strategy of model-based science. *Biol. Philos.* **2006**, *21*, 725–740. [[CrossRef](#)]
47. Harrison, A.G.; Treagust, D.F. Learning about atoms, molecules, and chemical bonds: A case study of multiple-model use in grade 11 chemistry. *Sci. Educ.* **2000**, *84*, 352–381. [[CrossRef](#)]
48. Malone, K.L.; Schuchardt, A.; Sabree, Z. Models and modeling in evolution. In *Evolution Education Re-Considered*; Harms, U., Reiss, M.J., Eds.; Expected Publication; Springer: Berlin, Germany, in press.
49. Malone, K. Correlations among knowledge structures, force concept inventory, and problem-solving behaviors. *Phys. Rev. Spec. Top. Phys. Educ. Res.* **2008**, *4*, 020107. [[CrossRef](#)]
50. Jackson, J.; Dukerich, L.; Hestenes, D. Modeling Instruction: An effective model for science education. *Sci. Educ.* **2008**, *17*, 10–17.
51. Liang, L.L.; Fulmer, G.W.; Majerich, D.M.; Clevinstine, R.; Howanski, R. The effects of a model-based physics curriculum program with a physics first approach: A causal-comparative study. *J. Sci. Educ. Technol.* **2012**, *21*, 114–124. [[CrossRef](#)]
52. Malone, K.L.; Schuchardt, A. The efficacy of Modeling Instruction in chemistry: A case study. In Proceedings of the HICE 2016: The 14th Annual Hawaii International Conference on Education, Honolulu, HI, USA, 3–6 January 2016; pp. 1513–1518.
53. Schuchardt, A.; Malone, K.L.; Diehl, W.; Harless, K.; McGinnis, R.; Parr, T. A case study of student performance following a switch to a modeling-based physics first course sequence. In Proceedings of the NARST International Conference, Baltimore, MD, USA, 30 March–2 April 2008.
54. Momsen, J.L.; Long, T.M.; Wyse, S.A.; Ebert-May, D. Just the facts? Introductory undergraduate biology courses focus on low-level cognitive skills. *CBE-Life Sci. Educ.* **2010**, *9*, 435–440. [[CrossRef](#)] [[PubMed](#)]
55. Adamson, S.L.; Banks, D.; Burtch, M.; Cox, F.I.I.; Judson, E.; Turley, J.B.; Benford, R.; Lawson, A.E. Reformed undergraduate instruction and its subsequent impact on secondary school teaching practice and student achievement. *J. Res. Sci. Teach.* **2003**, *40*, 939–957. [[CrossRef](#)]
56. Koenig, K.; Schen, M.; Bao, L. Explicitly targeting pre-service teacher scientific reasoning abilities and understanding of nature of science through an introductory science course. *Sci. Educ.* **2012**, *21*, 1.

57. Moore, N.; O'Donnell, J.; Poirier, D. Using cognitive acceleration materials to develop pre-service teachers' reasoning and pedagogical Expertise. In *Advancing the STEM Agenda in Education*; Madison University of Wisconsin-Stout: Menomonie, WI, USA, 2012.
58. Koba, S.; Wojnowski, B. *Exemplary Science: Best Practices in Professional Development*, 2nd ed.; NSTA Press: Arlington, VA, USA, 2013; ISBN 978-1-936959-07-5.
59. Loucks-Horsley, S.; Stiles, K.E.; Mundry, S.; Love, N.; Hewsen, P.W. *Designing Professional Development for Teachers of Science and Mathematics*, 3rd ed.; Corwin Press: Thousand Oaks, CA, USA, 2010; ISBN 1412974143.
60. Hewson, P. Teacher professional development in science. In *Handbook of Research in Science Education*; Abell, S.K., Lederman, H.J., Eds.; Lawrence Erlbaum Associates: Mahwah, NJ, USA, 2007; pp. 1179–1203.
61. Darling-Hammond, L.; Chung Wei, R.; Andree, A.; Richardson, N.; Orphanos, S. *Professional Learning in the Learning Profession: A Status Report on Teacher Development in the United States and Abroad*; National Staff Development Council: Oxford, OH, USA, 2009.
62. Garet, M.S.; Porter, A.C.; Desimone, L.; Birman, B.F.; Kwang, S.Y. What makes professional development effective? Results from a national sample of teachers. *Am. Educ. Res. J.* **2001**, *38*, 915–945. [CrossRef]
63. Weiss, I.R.; Pasley, J.D. Scaling Up Instructional Improvement through Teacher Professional Development: Insights from the Local Systemic Change Initiative. Consortium for Policy Research in Education (CPRE) Policy Briefs. 2006. Available online: [http://repository.upenn.edu/cpre\\_policybriefs/32](http://repository.upenn.edu/cpre_policybriefs/32) (accessed on 12 June 2018).
64. Supovitz, J.A.; Turner, H.M. The effects of professional development on science teaching practices and classroom culture. *J. Res. Sci. Teach.* **2000**, *37*, 963–980. [CrossRef]
65. Rhoton, J.; Wojnowski, B. Building ongoing and sustained professional development. In *Teaching Science in the 21st Century*; Shane, P., Rhoton, J., Eds.; National Science Teachers Association and the National Science Education Leadership Association: Washington, DC, USA, 2005; pp. 113–126.
66. Shane, P.; Wojnowski, B. Technology integration enhancing science: Things take time revisited. *Sci. Educ.* **2007**, *16*, 51–57.
67. Van Driel, J.H.; Beijaard, D.; Verloop, N. Professional development and reform in science education: The role of teachers' practical knowledge. *J. Res. Sci. Teach.* **2001**, *38*, 137–158. [CrossRef]
68. Inhelder, B.; Piaget, J. *The Growth of Logical Thinking from Childhood to Adolescence: An Essay on the Construction of Formal operational Structures*; Routledge & Kegan Paul: London, UK, 1958; ISBN 9780415210027.
69. Abd-El-Khalick, F.; Lederman, N.G. Improving science teachers' conceptions of nature of science: A critical review of the literature. *Int. J. Sci. Educ.* **2000**, *22*, 665–701. [CrossRef]



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).