
International Journal of Science Education
Publication details, including instructions for authors and subscription information:
http://www.informaworld.com/smpp/title~content=t713737283

# Examining the Relationship Between Students' Understanding of the Nature of Models and Conceptual Learning in Biology, Physics, and Chemistry

Janice D. Gobert[a]; Laura O'Dwyer[b]; Paul Horwitz[c]; Barbara C. Buckley[d]; Sharona Tal Levy[e]; Uri Wilensky[f]

[a] Department of Social Sciences, Worcester Polytechnic Institute (WPI), Massachusetts, USA [b] Department of Educational Research, Measurement and Evaluation, Boston College, Massachusetts, USA [c] The Concord Consortium, Massachusetts, USA [d] Mathematics, Science, and Technology Program, WestEd, California, USA [e] Faculty of Education, University of Haifa, Haifa, Israel [f] Learning Sciences, Northwestern University, Illinois, USA

First published on: 14 June 2010

**To cite this Article** Gobert, Janice D. , O'Dwyer, Laura , Horwitz, Paul , Buckley, Barbara C. , Tal Levy, Sharona and Wilensky, Uri(2010) 'Examining the Relationship Between Students' Understanding of the Nature of Models and Conceptual Learning in Biology, Physics, and Chemistry', International Journal of Science Education,, First published on: 14 June 2010 (iFirst)

**To link to this Article: DOI:** 10.1080/09500691003720671

**URL:** http://dx.doi.org/10.1080/09500691003720671

*International Journal of Science Education*
2010, 1–32, iFirst Article



**RESEARCH REPORT**

# Examining the Relationship Between Students' Understanding of the Nature of Models and Conceptual Learning in Biology, Physics, and Chemistry

Janice D. Gobert[a]*, Laura O'Dwyer[b], Paul Horwitz[c], Barbara C. Buckley[d], Sharona Tal Levy[e] and Uri Wilensky[f]
[a]*Department of Social Sciences, Worcester Polytechnic Institute (WPI), Massachusetts, USA;* [b]*Department of Educational Research, Measurement and Evaluation, Boston College, Massachusetts, USA;* [c]*The Concord Consortium, Massachusetts, USA;* [d]*Mathematics, Science, and Technology Program, WestEd, California, USA;* [e]*Faculty of Education, University of Haifa, Haifa, Israel;* [f]*Learning Sciences, Northwestern University, Illinois, USA*



This research addresses high school students' understandings of the nature of models, and their interaction with model-based software in three science domains, namely, biology, physics, and chemistry. Data from 736 high school students' understandings of models were collected using the Students' Understanding of Models in Science (SUMS) survey as part of a large-scale, longitudinal study in the context of technology-based curricular units in each of the three science domains. The results of ANOVA and regression analyses showed that there were differences in students' pre-test understandings of models across the three domains, and that higher post-test scores were associated with having engaged in a greater number of curricular activities, but only in the chemistry domain. The analyses also showed that the relationships between the pre-test understanding of models subscales scores and post-test content knowledge varied across domains. Some implications are discussed with regard to how students' understanding of the nature of models can be promoted.

Keywords: *Inquiry-based teaching; Learning environment; Model-based learning; Nature of science; Science education; Nature of models*

*Corresponding author. Department of Social Sciences, Worcester Polytechnic Institute (WPI), Atwater-Kent, Room 103, Worcester, MA 01609, USA. Email: jgobert@wpi.edu

ISSN 0950-0693 (print)/ISSN 1464-5289 (online)/10/000001–32

DOI: 10.1080/09500691003720671

## Introduction

The largest thrust in science education research has traditionally been to characterize and promote students' science learning with a focus on changes in content knowledge and the conceptual change processes that lead to this change (cf. Driver, Guesne, & Tiberghien, 1985; Driver, Leach, Millar, & Scott, 1996; Driver, Squires, Rushworth, & Wood-Robinson, 1994; Posner, Strike, Hewson, & Gertzog, 1982; Strike & Posner, 1985). More recently, however, there has been an acknowledgment within science education that learners' understanding of the nature of science has a significant impact on students' science learning itself (Carey & Smith, 1993, 1995; Hammer, 1994, 1995; Hammer & Elby, 2002; Linn, Songer, Lewis, & Stern, 1991; Perkins, Jay, & Tishman, 1993; Songer & Linn, 1991). Relevant to this is the finding that, in addition to content misconceptions (cf. Clement, Brown, & Zietsman, 1989), students come to science instruction with naïve theories and/or misconceptions about the nature of science (Driver et al., 1996; Grosslight, Unger, Jay, & Smith, 1991) and that these beliefs about science impact students' understanding of the content knowledge. Further, it has been suggested that students must make changes to these naïve understandings of the nature of science in order to more deeply understand both domain-specific theories in science and content itself (Champagne, Gunstone, & Klopfer, 1985; Snir, Smith, & Grosslight, 1988).

In a similar vein to the research above is the perspective that students' understanding of the nature of scientific models is also critical to their understanding of science content (Gobert & Discenna, 1997; Gobert & Pallant, 2004; Schwarz, 2002; Schwarz & White, 2005). Additionally, we conceptualize students' understanding of models as a component or subset of their understanding of the nature of science (Gobert & Pallant, 2004; Schwarz, 2002; Schwarz & White, 2005). This will be discussed later in the paper.

The research described here examines the nature of students' understanding of the nature of models, the impact of model-based curricula on changes to these understandings, and the relationship between students' understanding of the nature of models and their learning in the domains of physics, chemistry, and biology.



## Overview of the Research Literature

### *How do Students' Understandings of the Nature of Science 'in General' Interact with Learning?*

There are a fair number of studies that have examined the relationship between students' understanding of the nature of science and its relationship to content learning (Carey & Smith, 1993, 1995; Hammer, 1994, 1995; Linn et al., 1991; Perkins et al., 1993; Perry, 1970; Songer & Linn, 1991). For example, using correlational techniques, Hammer (1994, 1995) and Songer and Linn (1991) both showed that more sophisticated understanding of the nature of science, i.e., that it is a dynamic enterprise in which evidence changes over time, may contribute to better learning of science content. Specifically, Hammer (1994) found that students who believed that scientific knowledge was coherent also tended to be more careful about building an integrated conceptual understanding. In the study by Songer and Linn (1991), students who believed that science was relevant to everyday problems were more likely to seek to understand underlying scientific principles and apply them to new situations. Other studies have investigated the influence of science learning on views of science (Carey & Smith, 1993; Chen & Klahr, 1999); however, results here were not consistently positive (Abd-El-Khalick & Lederman, 2000; Burbules & Linn, 1991).

In addition to the correlational studies described above, intervention studies have been conducted with the goal of promoting students' understanding of the nature of science. Among these there has been some success at moving students' understanding of the nature of science along the spectrum from a naïve, i.e., science is a collection of facts, to a more sophisticated understanding of science, i.e., it is a complex body of knowledge that changes as empirical findings influence it. Carey, Evans, Honda, Jay, and Unger (1989) tested whether students could articulate more sophisticated understandings of science following an innovative science curriculum. They found that students made progress in differentiating between data and hypotheses and could also see how ideas tested were limited in ways in which experiments were conducted. Hennessey and Beeth (1993) and Hennessey (1995) also reported gains regarding students' understanding of the nature of science as a result of a curricular intervention. Bell (1998) scaffolded students to use data in order to substantiate their arguments in a debate task about a controversial science concept; he found that students made gains in their understanding of the nature of science as well as significant gains in content understanding.



### *How do Students' Understandings of the Nature of Models 'Specifically' Influence Science Learning?*

Previous research has shown that students possess little knowledge about the nature and purpose of scientific models (Carey & Smith, 1993; Schwarz & White, 1999; van Driel & Verloop, 1999). Although it is difficult to empirically disentangle modeling knowledge from content knowledge (Schwarz, 2002), some studies have shown that a learner's understanding of models is significantly related to students' science learning (Gobert & Discenna, 1997; Gobert & Pallant, 2004; Schwarz, 2002; Schwarz & White, 2005). Regarding correlational studies, Smith, Maclin, Houghton, and Hennessy (2000) found that students who understood that models can be used as explanatory tools also used models to explain evidence. Gobert and Discenna (1997) found that students who held a more sophisticated understanding of the nature of models, i.e., that models are tools for scientific reasoning, were better able to make inferences with their models once constructed. A very recent study (Sins, Savelsbergh, van Joolingen, & van Hout-Wolters, 2009) studied the relationship between students' understanding of models and the depth of cognitive processing as measured by using think-aloud protocols. Results showed a positive correlation between students' understanding of models and *deeper* processing of material presented, as well as a negative correlation between students' understanding of models and *shallow* processing of material presented.

In terms of intervention studies designed to promote students' understandings of models, Schwarz and White (2005) showed that the METT (Model-enhanced ThinkerTools) curriculum, which was designed to teach about the nature of models, was successful at clarifying and broadening students' understanding of the nature and purpose of models, as well as inquiry skills and physics knowledge. Honda (1994) found that 11th-grade students achieved gains in understanding the nature of science after a brief curriculum unit that specifically addressed modeling. Gobert and colleagues (Gobert & Pallant, 2004; Gobert, Slotta, & Pallant, 2002; Gobert, Snyder, & Houghton, 2002) wrote a curriculum that engaged students in many model-based inquiry tasks with some explicit instruction in the nature of models. This curriculum yielded significant gains in students' understanding of the nature of models as evidenced in their students' written responses to open-ended questions about the nature and purpose of models. Lastly, Wilensky and colleagues studied differences between students' types of scientific models and found that students' understanding of scientific models contained more causal and mechanistic elements when they worked with computer-based multi-agent models when compared to equation-based models (Sengupta & Wilensky, 2009; Wilensky, 1999b; Wilensky & Reisman, 2006). In sum, results such as these suggest that fostering students' understandings of models is possible, particularly when students are engaged in model-based activities that are rich and properly scaffolded. What these studies have not shown is *whether* more sophisticated understandings of the nature of models, i.e., models as tools with which to reason and experiment, influence students' content learning when learning with models. This is one of the research questions addressed here.



### *The Need for Modeling, and the Nature of Models as a Subset of Nature of Science*

Understanding models is an important aspect of one's understanding of science, since models and modeling play such a large role in scientific discovery and science learning at all levels of education. Others concur that modeling can have a significant impact on lifelong learning and scientific literacy (Bisard, Aron, Francek, & Nelson, 1994; Linn & Muilenberg, 1996). Models are generative and as such afford more flexible knowledge use and transfer to other science concepts; thus, model-based pedagogical approaches have the potential to impact scientific literacy more than do traditional curricula (Gobert & Horwitz, 2002). Chittleborough, Treagust, Mamiala, and Mocerino (2005) point out that the use of models requires the learner to identify the analogue and without that connection the model has no value. Chittleborough et al.'s (2005) work is an important first step in recognizing the role of the learner in recognizing the relationship between the model and the scientific object/processes it represents. We too acknowledge the need on the part of the students to map between the object/process and its model. Furthermore, underlying our work is the belief that the efficacy of models as representational tools for deep learning rests, at least in part, on students' understanding of models as abstracted representations of scientific phenomena; others concur with this claim (Justi & Gilbert, 2002a; Treagust, Chittleborough, & Mamiala, 2002).

The utility of and need for models and modeling tasks in science instruction has been broadly acknowledged (Clement, 1993; Giere, 1990; Gilbert, 1993; Gobert & Discenna, 1997; Hesse, 1963; Linn & Muilenberg, 1996; NRC, 1996), and research in order to unpack the relationship between students' understanding of models and its relationship to science learning is of critical importance (Schwarz & White, 2005).

As previously stated, students' understanding of the nature of scientific models is critical to both their understanding of the nature of science (Gobert & Discenna, 1997) and their understanding of science content (Gobert & Discenna, 1997; Gobert & Pallant, 2004; Schwarz, 2002; Schwarz & White, 2005; Smith et al., 2000; Treagust et al., 2002). In this paper, our presupposition is that the understanding of scientific models is an important component of students' understandings of the nature of science as a whole (cf. Lederman, 1992, 2006). Others concur with this approach (Schwarz, 2002; Schwarz & White, 1999, 2005; Wilensky & Reisman, 2006).

Briefly, the key connection between the nature of models and the nature of science relates to the belief that models are to be viewed as not completely accurate from a scientific point of view; that is, they are tentative and open to further revision and development (Crawford & Cullin, 2004). Additionally, a key concept is that there can be multiple models for the same scientific object/process and that the model put forth can depend on the perspective of the scientist and the purpose of the research being conducted. Lastly, a model is a tool for other scientists to discuss, debate, etc. (Sins et al., 2009). As such, the process or nature of science can be thought of as an endeavor by which competing models are developed, tested, and compared (Giere, 1990; Hestenes, 1987; Justi & Gilbert, 2002b).

Viewing the nature of models as a subset of the nature of science is compatible with current views of scientific literacy. Specifically, Hodson (1992) has characterized the purpose of science *education* as: (1) *the learning of science*, i.e., to understand the ideas produced by science; (2) *learning about science*, i.e., to understand important issues in the philosophy, history, and methodology of science; and (3) *learning to do science*, i.e., to be able to take part in those activities that lead to the acquisition of scientific knowledge. These three purposes, in particular the second and third, propose that models and modeling activities must play a central role in the methodology of science (2 above), and in learning how to do science that helps produce scientific knowledge (3 above, see Justi & Gilbert, 2002a). Others have made similar claims (Schwarz, Meyer, & Sharma, 2007). Specific to the argument in this paper is the claim that *learning about science* (Hodson's second component of science *education*) means that students should have an understanding of *how scientific knowledge is generated;* to us, this means that students should understand the role that models play in the accreditation and dissemination of the products of scientific inquiry (Justi & Gilbert, 2002a).

## Rationale, Context, and Sample

### *Rationale*

As a way to begin to unpack the relationship between students' understanding of the nature of models and their science learning, we address this relationship in three different science domains, namely, biology (genetics), physics (Newtonian mechanics), and chemistry (gas laws).

This line of research is important for a number of reasons. First, this study can provide data about the contextualized nature of learning with models within each of these domains (Hofer, 2000). Second, this research can provide information about students' understanding of models and their relationship to content learning. At present, little is known about how a sophisticated understanding of models, i.e., that models are representations to test theories, reason with, etc., may affect learning with models (Schwarz, 2002). Finally, this study can provide insights about the design and refinement of instructional strategies that promote students' understanding of models; this is important because the connection between research in scientific epistemological understandings and curriculum design is not well understood (Smith & Wenk, 2003).

### *Context and Curricular Materials*

To examine the nature of students' understanding of models, the impact of model-based curricula on changes in understanding, and the relationship between students' understanding of models and learning in physics, chemistry, and biology, data were collected as part of the Modeling Across the Curriculum (MAC) project. The MAC project was a large-scale project that was funded by the Interagency Education Research Initiative (IERI) program (IERI no. 0115699, http://mac.concord.org). One of the goals of the MAC project was to examine whether there were measurable learning gains across grades from exploration and inquiry of curricula based on computer models of core science content. To address this goal, we conducted both longitudinal and cross-sectional research to determine the effects of engaging students in model-based activities across multiple years and multiple domains of science. All the students in the study were in high school, and the grade levels ranged from the 9th through the 12th grade. We measured gains in content knowledge in each domain using computer-scored, multiple-choice instruments of our own design. We assessed students' understandings of models separately for each content domain, using the students' Understanding of Models in Science (SUMS) survey (Treagust et al., 2002). The research described here uses cross-sectional data collected during the 2005–2006 school year from high school students.

The curricular materials for the MAC project were in the form of interactive activities dealing with scientific processes and phenomena and were based on computer models of the relevant scientific domain. Each activity behaves according to the laws and rules of that domain. For example, the biology curriculum, BioLogica™, embodies Mendel's famous laws of genetics and other models of inheritance so that changes made to an organism's genotype result in phenotypic changes, as appropriate, and crosses between organisms produce the correct statistical distribution of offspring genotypes. The physics curriculum, Dynamica™, is a model of Newtonian mechanics as it applies to point particles, and the chemistry curriculum, Connected Chemistry[1] (Levy, Kim, & Wilensky, 2004; Levy & Wilensky, 2009; Stieff & Wilensky, 2002, 2003), approaches learning about the gaseous phase using multi-agent NetLogo models (Wilensky, 1999a) that highlight the system's emergent nature. As students work through our curricular packages, they are presented with inquiry tasks similar to those described by the National Science Education Standards (NRC, 1996). The inquiry tasks include making predictions with representations and models, designing and conducting experiments with models, interpreting data from representations (e.g., models, graphs, Punnett squares) and experiments, generating explanations using models, and generating and using equations to represent the behavior of models. The curricular packages, each about two to three weeks worth of classroom activities, are described briefly as follows:



- BioLogica™: The MAC biology curriculum consists of 12 activities that teach genetics through increasingly elaborate models of the parts, processes, and mechanisms relevant to that domain. The model includes Mendelian inheritance plus incomplete dominance, sex linkage, and polygenicity, as well as meiosis and mutations.
- Dynamica™: The MAC physics curriculum consists of nine activities for teaching and exploring the effect of forces on point masses. Using a set of objects that includes masses, forces, walls, and targets, Dynamica's real-world analogs range from billiard balls to rocket ships. The units cover vectors, graphs, forces in one and two dimensions, and motion in a uniform gravitational field.
- Connected Chemistry: The MAC chemistry curriculum consists of seven activities that address learning about the gaseous phase, the gas laws, and kinetic molecular theory. Topics include the effects of temperature, volume, and the number of particles on the pressure exerted by a contained gas and construction of the gas law equations. The curriculum emphasizes how microscopic particles' properties and interactions result in the emergence of macroscopic phenomena.

In the design of our learning activities for each of these curricula, we used a progressive model-building approach in which simpler models provide conceptual leverage for more complex models. This approach has demonstrated success at promoting deep conceptual understanding in science (Gobert & Clement, 1999; Raghavan & Glaser, 1995; White & Frederiksen, 1990). In addition, we draw on literature about students' learning with diagrams (cf. Gobert, 2000; Gobert & Clement, 1999; Larkin & Simon, 1987; Lowe, 1993), model-based learning (Gobert & Buckley, 2000), and students' learning difficulties with models (Lowe, 1989) to inform our scaffolding design; thus, in summary our scaffolding was designed to support students' progressive model-building so as to foster deeper learning of the content in each of the domains. The types of scaffolding we designed and used in these activities are (Gobert, Buckley, & Clarke, 2004):

- *Representational assistance* to guide students' understanding of the representations or the domain-specific conventions in the domain and to support students in using multiple representations.
- *Model pieces acquisition* to focus students' attention on the perceptual pieces of the representations and support students' knowledge acquisition about one or more aspects of the phenomenon (e.g., spatial, causal, functional, temporal).
- *Model pieces integration* to help students combine model components in order to come to a deeper understanding of how they work together as a causal system.
- *Model-based reasoning* to support students' reasoning with models, i.e., inference-making, predictions, and explanations.
- *Reconstruct, reify, and reflect* to support students to refer back to what they have learned, reinforce it, and then reflect to move to a deeper level of understanding.

It is important to note that we did not explicitly scaffold students' understandings of models in either the physics or biology curricula. However, for the chemistry domain, the Connected Chemistry curriculum weaves in a distinct strand of instruction addressing modeling issues. For example, the students are asked to construct theoretical models, examine the models' rules, and compare the model with the phenomena it represents (see Wilensky, 2001; Wilensky & Reisman, 2006, for a fuller account of this approach).



### *Teacher's Role*

For the MAC project, we did not directly prescribe how teachers should use our learning activities. In our partner schools, with whom we worked more closely, we asked teachers to use the activities in the sequence provided, but we did not specify whether the activities were to be used to introduce, experience, or review a concept, or as some combination thereof. Even with that request, some teachers chose not to use all of the activities. For this reason, we used our log files, generated automatically as learners' used the software, to document the usage for each individual student. This permitted us to know the number of core and optional activities engaged in by each student; these data were used in our analyses and will be described in more detail later.

### *Research Questions and Sample*

Using data collected as part of the MAC project, this paper specifically addresses the following research questions:

(1) What are students' understandings of models in the domains of physics, chemistry, and biology at the outset of the study (as measured by pre-test assessments)?
(2) Does engaging in modeling enhance students' understanding of the nature of models?
(3) Do those students with a more sophisticated understanding of the nature and purpose of models, i.e., that models are tools to test hypotheses, reason with,

etc., at the outset of the study have higher post-test content scores in each of the three domains? That is, is one's understanding of the nature of models a significant predictor of content learning?

These research questions allow us to better unpack the relationship between the nature of students' understanding of the nature of models and their role in learning in each domain.

Participating students were drawn from 13 high schools from across the USA whose principal or science department head volunteered to participate in our project; thus, individual students did not themselves volunteer to participate. We refer to our participating schools as our partner schools. These schools represent a wide range of socio-economic levels (estimated by the percentage of students in the school who receive free or reduced lunch) with schools ranging from 0% to 41% of students receiving free or reduced lunch. Across all 13 schools, the average percentage of students receiving free or reduced lunch was 16%. School size ranged from 120 students to slightly more than 2,000 students, with the average school size 1,200 students. The average number of students in the science classes in these schools was 24 students. The data used to address the research questions were drawn from students who:

(1) had completed a sequence of MAC activities in a particular scientific domain during the school year 2005–2006;
(2) had completed the pre- and post-SUMS (Treagust et al., 2002) in 2005–2006; and
(3) had not participated in any of our other domain interventions in prior years; thus, they completed the SUMS instrument in one domain only and had not taken either of our other two curricular packages.

Using these criteria, this research uses data from 420 physics students, 218 chemistry students, and 98 biology students from our 13 partner schools. Table 1 provides descriptive statistics regarding students' ages in each domain (later in the paper, we present statistical analyses of these data (Tables 5 and 6)).

Table 1. Mean age of participants by domain

| | | Descriptive statistics | | | | |
|---|---|---|---|---|---|---|
| Domain | | *N* | Minimum | Maximum | Mean | *SD* |
| Dynamica™ | age_at_admin | 418 | 14.00 | 18.00 | 15.9904 | 1.18135 |
| | Valid N (listwise) | 418 | | | | |
| Connected Chemistry | age_at_admin | 217 | 16.00 | 19.00 | 16.8249 | .62858 |
| | Valid N (listwise) | 217 | | | | |
| Biologica™ | age_at_admin | 98 | 14.00 | 18.00 | 15.5816 | .64093 |
| | Valid N (listwise) | 98 | | | | |

## Methods and Instrumentation

As previously stated, for this research we used data from students who were participating in one domain only from the MAC curricula for the first time. This ensured that we were not pooling data from students who had engaged in more than one of our modeling curricula or had completed any of the instruments in more than one year. For example, the reason for the relatively low number of biology students in our study is that most of the students who took biology in 2005–2006 had already taken physics or chemistry and thus these students had been exposed both to the survey and to one or two of our modeling curricula. For this reason, they were excluded from these analyses.

### *Content Knowledge Measures*

For each content area, identical pre- and post-test content measures were administered. The items are all multiple-choice questions which were designed to assess students' content knowledge in each of the domains with particular focus on targeting problematic areas as reported in the science education literature.

### *Students' Understanding of Models in Science (SUMS) Survey*

The SUMS survey was used to collect data about students' understanding about the nature of models in sciences (the complete survey is included in Table 2). The SUMS survey was administered online to student participants in the context of their science class in one of the three domains; the survey was administered both before and after our curricular materials were used in each of the content areas addressed.

The SUMS survey was developed by Treagust et al. (2002) based both on their earlier work (Treagust, Chittleborough, & Mamiala, 2001) and the earlier work of Grosslight et al. (1991), who designed an open-ended survey to assess students' understanding of models, namely, about models and the uses of models in science. The SUMS instrument made use of the Grosslight et al. items, but rather than present these using open-response format, they asked students to rate the items using a 1–5 Likert scale. The 26 items[2] are presented with a statement about the nature and role of models in science and are asked to endorse or oppose the statements on a scale ranging from 'strongly disagree' to 'strongly agree.' Using a Likert scale survey does pose some limitations regarding the richness of the data collected, but it affords the broad scalability of usage, important to this project.

In Treagust et al.'s (2002) research using the SUMS instrument, the authors administered the survey on 228 students in Grades 8, 9, and 10 from two schools in Australia. Analyses of the data collected revealed a five-factor solution that represented measures of five subscales or constructs relating to models. These were: (1) models as multiple representations (MR), (2) models as exact replicas (ER), (3) models as explanatory tools (ET), (4) uses of scientific models (USM),

Table 2. Modeling across the curriculum epistemological survey

| Item | MAC survey no. | SUMS survey no.[a] (construct name) |
|---|---|---|
| Models are used to: | | |
| Show an idea | 1 | 20 (ET) |
| Explain scientific phenomena | 2 | 19 (ET) |
| Physically or visually represent something | 3 | 17 (ET) |
| Show a smaller scale size of something | 4 | 16 (ER) |
| Show the relationship of ideas clearly | 5 | 3 (ER) |
| Make and test predictions about a scientific event | 6 | 24 (USM) |
| Help formulate ideas and theories about scientific events | 7 | 22 (USM) |
| Help create a picture in your mind of the scientific happening | 9 | 18 (ET) |
| A model may be changed if: | | |
| There are changes in data or beliefs | 10 | 27 (CNM) |
| There are new findings | 11 | 26 (CNM) |
| New theories or evidence prove otherwise | 12 | 25 (CNM) |
| A model needs to be close to the real thing by: | | |
| Being very exact, so nobody can disprove it | 13 | 11 (ER) |
| Being very exact in every way except for size | 14 | 13 (ER) |
| Being as close as it can be to the real thing | 15 | 10 (ER) |
| Giving the correct information and showing what the object/thing looks like | 16 | 14 (ER) |
| The features of a model are as follows: | | |
| It should be an exact replica | 17 | 9 (ER) |
| Everything about it should be able to tell what it represents | 18 | 12 (ER) |
| It shows what the real thing does and what it looks like | 19 | 15 (ER) |
| Has what is needed to show or explain a scientific phenomenon | 20 | 8 (MR) |
| It can be a diagram or a picture, a map, graph or a photo | 21 | 21 (ET) |
| Many models may be used to express features of a science phenomenon by showing: | | |
| Different versions of the phenomenon | 22 | 2 (MR) |
| Different sides or shapes of an object | 23 | 5 (MR) |
| Different parts of an object or showing the objects differently | 24 | 6 (MR) |
| Different perspectives to a view an object | 25 | 1 (MR) |
| How different information is used | 26 | 7 (MR) |
| How it depends on individuals different ideas on what things look like or how they work | 27 | 4 (MR) |

[a]Item 23 on the SUMS survey was excluded from the analyses as the wording of the item was considered problematic (*models are used to show how they are used in scientific investigations*).

and (5) the changing nature of models (CNM). Treagust et al. (2002) provide reliability data on each of the scales of the SUMS survey; the reliability of the scales ranged from 0.71 to 0.84, thus the instrument has high internal consistency for each scale; item-to-total correlations were above 0.45 with the exception of one item (Item 16).

The mean of students' scores across the individual items that comprised each subscale was calculated and used to represent students' score on each subscale. With the exception of the ER subscale, a higher scale score represented a more sophisticated understanding of the nature and role of models in science, as measured by a 1–5 Likert scale. On the ER subscale, students who endorsed the items more strongly (i.e., either 'agree' or 'strongly agree') were those who held a more naïve understanding of models in science, i.e., that they were like mini-replicas of the objects they represent. We adopted these five measurement scales in our research, comparing students' scores on each subscale before and after the MAC intervention in each of the respective content domains, and making comparisons across science domains.



Before conducting the analysis to address our research questions, we first sought to compare the reliabilities of the five measurement scales as reported by Treagust et al. (2002) to those found for our sample of students. As previously stated, the reliabilities for Treagust et al.'s scales were high, ranging from 0.71 to 0.84. Table 3 shows that the pre-test reliabilities of the ET and ER scales in the physics domain were lower than those reported for Treagust et al.'s data (0.69 for both scales). For the pre- and post-test scores in the biology domain, the reliabilities for the ET and USM scales were lower than those reported by Treagust et al. (2002). Mean scores for the five measurement scales across the three domains ranged from moderate to strong for both the pre- and post-tests. This indicates that the ways in which students understand scientific models and how they are used are moderately related to each other.

## Results

### *Analysis for Research Question 1: Students' understanding of the nature of models in each domain*

To examine the nature of students' understanding of models in each domain at the outset of the study, we examined the mean SUMS pre-test scores and compared them across the three domains.

Analysis of variance (ANOVA) was used to examine whether the differences in the pre-test means were statistically significant across domains, and Bonferroni post-hoc tests were used to isolate the specific differences between pairs of means. Table 4 presents the measurement subscale means and standard deviations for the pre-test SUMS administration in each domain. For three of the five measured constructs, the analyses revealed statistically significant differences between pre-test means across the domains, namely, ET ($F_{2,\,733} = 4.36$; $p < .05$), MR ($F_{2,\,733} = 3.12$; $p <$

Table 3. Reliabilities of the measurement scales

| | Reported reliability (Treagust et al., 2002) | Physics ($N$ = 420) | | Chemistry ($N$ = 218) | | Biology[a] ($N$ = 98) | |
|---|---|---|---|---|---|---|---|
| | | Pre-test | Post-test | Pre-test | Post-test | Pre-test | Post-test |
| Models as explanatory tools (ET) | 0.81 | 0.69 | 0.81 | 0.76 | 0.86 | 0.59 | 0.65 |
| Models as exact replicas (ER) | 0.84 | 0.69 | 0.80 | 0.73 | 0.76 | 0.72 | 0.70 |
| Models as multiple representations (MR) | 0.71 | 0.80 | 0.81 | 0.80 | 0.84 | 0.76 | 0.76 |
| Uses of scientific models (USM) | 0.72 | 0.74 | 0.76 | 0.67 | 0.75 | 0.56 | 0.62 |
| Changing nature of models (CNM) | 0.73 | 0.79 | 0.84 | 0.81 | 0.90 | 0.73 | 0.81 |

[a]One item was removed from the MR construct for the biology domain because no students in the sample responded to this survey item.

.05), and USM ($F_{2,\,733} = 3.65$; $p < .01$). For these three constructs, post-hoc tests revealed significant differences between the pre-test means in physics and biology; additionally, for the USM subscale, significant differences were observed between chemistry and biology as well as between physics and biology. There were no statistically significant differences found for the subscales ER or CNM across any of the three domains.

These results show that the students who participated in the three domains were not entirely similar with respect to their pre-intervention understanding of models. Specifically, students who participated in the physics curriculum appear to have had a more sophisticated understanding of ET, MR, and the USM when they began the physics curricula than did the students who participated in the biology curriculum. Additionally, students who participated in the chemistry curriculum appear to have had a more sophisticated understanding of the USM when they began the curriculum when compared to the students who participated in the biology curriculum. Thus, in general, the students who were about to partake in the biology curriculum had a more naïve understanding of models in science, as reflected by: (1) lower scores on the 1–5 Likert subscales for ET and MR when compared to those in physics, and (2) lower scores on the 1–5 Likert subscale for USM when compared to those in chemistry.

These differences prompted us to address whether students in the biology curriculum might be younger in age, on average, when compared to the cohorts in each of the physics and chemistry groups. An ANOVA was conducted in order to test this. As can be seen in Table 5, the three groups are statistically different from each other in terms of age ($F_{2,\,703} = 72.442$; $p < .001$). Post-hoc analyses, as shown in Table 6, revealed that biology students were the youngest of the three groups and chemistry students were the oldest. Chemistry students were significantly older than physics students ($p < .001$). Chemistry students were significantly older than biology students ($p < .001$). The smallest difference in age was between biology and physics students; physics students were slightly older than the biology students and the difference was significant ($p < .01$). Later in the 'Discussion' section we address these findings regarding age in terms of our original research question.

### *Analysis for Research Question 2: The effects of model-based learning on students' understanding of models*

To address our second research question, namely, whether engaging students in rich, authentic, model-based learning would enhance their understanding of the nature of models, we conducted regression analyses. Here we used the percentage of core curricular activities (other activities were designed to be optional or extension activities) that were used by students within a domain to predict their post-test SUMS scores on each of the five model subscales. These analyses were conducted using regression as opposed to ANOVA (comparing pre- and post-test means on each of the subscales in each of the science domains) because students differed in

Table 4. Differences between pre-test measurement subscale means across domains

| | Physics ($N = 420$) | | Chemistry ($N = 218$) | | Biology[a] ($N = 98$) | | ANOVA results | |
|---|---|---|---|---|---|---|---|---|
| | Mean (*SE*) | *SD* | Mean (*SE*) | *SD* | Mean (*SE*) | *SD* | *F* (df) | Significant |
| Pre-test ET score | 4.24 (0.02)[b] | 0.51 | 4.16 (0.04) | 0.57 | 4.07 (0.06)[b] | 0.55 | 4.36 (2, 733) | **.013** |
| Pre-test ER score | 3.72 (0.03) | 0.56 | 3.73 (0.04) | 0.55 | 3.79 (0.06) | 0.57 | 0.711 (2, 733) | .492 |
| Pre-test MR score | 3.87 (0.03)[b] | 0.51 | 3.86 (0.04) | 0.56 | 3.72 (0.06)[b] | 0.57 | 3.12 (2, 733) | **.045** |
| Pre-test USM score | 3.59 (0.04)[b] | 0.78 | 3.61 (0.05)[c] | 0.74 | 3.30 (0.07)[b,c] | 0.70 | 3.65 (2, 733) | **.002** |
| Pre-test CNM score | 4.19 (0.03) | 0.62 | 4.18 (0.05) | 0.74 | 4.06 (0.07) | 0.67 | 1.56 (2, 725) | .212 |

[a]One item was removed from the MR construct for the biology domain because no students in the sample responded to this survey item;
[b]Tukey's post-hoc tests indicate that the difference between pre-test means is statistically significant for physics and biology for $p$ is at least .05 (two-tailed);
[c]Tukey's post-hoc tests indicate that the difference between pre-test means is statistically significant for chemistry and biology for $p$ is at least .05 (two-tailed).
*Note*. Bold values indicate statistical significance for at least $p < .05$ (two-tailed)

Table 5. Analysis of variance of mean ages by domain

| age_at_admin | Sum of squares | df | Mean square | *F* | Significant |
|---|---|---|---|---|---|
| | ANOVA | | | | |
| Between groups | 140.349 | 2 | 70.175 | 72.442 | .000 |
| Within groups | 707.154 | 730 | .969 | | |
| Total | 847.503 | 732 | | | |

Table 6. Post-hoc comparison of mean age of participants by domain

| (*I*) domain | (*J*) domain | Mean difference (*I–J*) | *SE* | Significant | 95% confidence interval: Lower bound | 95% confidence interval: Upper bound |
|---|---|---|---|---|---|---|
| | Multiple comparisons (Bonferroni) | | | | | |
| Dynamica™ | Connected Chemistry | –.83445* | 0.08235 | .000 | –1.0321 | –.6369 |
| | BioLogica™ | .40880* | 0.11046 | .001 | .1437 | .6739 |
| Connected Chemistry | Dynamica™ | .83445* | 0.08235 | .000 | .6369 | 1.0321 |
| | BioLogica™ | 1.24325* | 0.11979 | .000 | .9558 | 1.5307 |
| BioLogica™ | Dynamica™ | –.40880* | 0.11046 | .001 | –.6739 | –.1437 |
| | Connected Chemistry | –1.24325* | 0.11979 | .000 | –1.5307 | –.9558 |

*Mean difference is significant at the 0.05 level.



terms of the number of modeling activities they engaged in both within and across the three different curricular packages. In these models, students' pre-test SUMS scores for each subscale were included as covariates.

Not surprisingly, students' pre-test SUMS scores were significant and positive predictors of their post-test SUMS scores on the five subscales and across the three domains with one exception in the biology domain (see Table 7). The exception was the regression model to predict students' post-test scores on the ER subscale in biology: students' pre-test subscale score was not a significant predictor of their post-test score on the same subscale ($\beta = 0.06$, $p = 0.537$).

Specific to this research question, the results of the regression analyses also show that for physics and biology, the percentage of core activities variable was *not* a significant predictor of students' post-test SUMS scores in any of the subscales (after controlling for students' pre-test SUMS scores). In contrast, however, for chemistry, the percentage of core activities variable was a weak but statistically significant predictor of students' post-test measures for the ET, MR, the USM, and the CNM subscales. Thus, in general, higher scores on these four subscales about

Table 7. Predicting the relationship between modeling and students' understanding of the nature of models

| | Physics ($N$ = 420) | | | Chemistry ($N$ = 218) | | | Biology[a] ($N$ = 98) | | |
|---|---|---|---|---|---|---|---|---|---|
| Outcome variable: post-test epistemology scores for each of the subscales | $\beta$ | *SE* | Significant | $\beta$ | *SE* | Significant | $\beta$ | *SE* | Significant |
| Constant | 2.10 | 0.28 | | 0.42 | 0.32 | | 2.02 | 0.45 | |
| ET pre-test score | **0.41** | **0.06** | **<.001** | **0.55** | **0.08** | **<.001** | **0.52** | **0.11** | **<.001** |
| Percent core activities | –0.01 | 0.17 | .783 | **0.17** | **0.21** | **<.01** | –0.19 | 0.33 | .050 |
| Percent of variance explained—total | 16.20 | | | 37.70 | | | 19.90 | | |
| Percent of variance explained—percent core activities | 0 | | | 2.2 | | | 0 | | |
| Constant | 1.76 | 0.25 | | 1.41 | 0.21 | | 3.38 | 0.47 | |
| ER pre-test score | **0.47** | **0.06** | **<.001** | **0.48** | **0.07** | **<.001** | 0.06 | 0.11 | .537 |
| Percent core activities | –0.05 | 0.17 | .337 | 0.10 | 0.18 | .103 | 0.01 | 0.33 | .898 |
| Percent of variance explained—total | 22.20 | | | 24.40 | | | 0 | | |
| Percent of variance explained—percent core activities | 0 | | | 0 | | | 0 | | |
| Constant | 2.37 | 0.24 | | 1.22 | 0.24 | | 1.71 | 0.38 | |
| MR pre-test score | **0.33** | **0.06** | **<.001** | **0.54** | **0.06** | **<.001** | **0.52** | **0.09** | **<.001** |
| Percent core activities | 0.07 | 0.16 | .160 | **0.19** | **0.16** | **<.01** | 0.08 | 0.26 | .366 |
| Percent of variance explained—total | 11.60 | | | 36.30 | | | 26.30 | | |
| Percent of variance explained—percent core activities | 0 | | | 3.10 | | | 0 | | |
| Constant | 2.31 | 0.23 | | 1.80 | 0.29 | | 2.35 | 0.51 | |
| USM pre-test score | **0.38** | **0.05** | **<.001** | **0.42** | **0.06** | **<.001** | **0.22** | **0.11** | **<.05** |
| Percent core activities | 0.07 | 0.20 | .174 | **0.17** | **0.23** | **<.01** | 0.11 | 0.42 | .295 |
| Percent of variance explained—total | 14.50 | | | 19.90 | | | 3.50 | | |
| Percent of variance explained—percent core activities | 0 | | | 2.40 | | | 0 | | |

Table 7. (*Continued*)

| Outcome variable: post-test epistemology scores for each of the subscales | Physics ($N = 420$) | | | Chemistry ($N = 218$) | | | Biology[a] ($N = 98$) | | |
|---|---|---|---|---|---|---|---|---|---|
| | $\beta$ | *SE* | Significant | $\beta$ | *SE* | Significant | $\beta$ | *SE* | Significant |
| Constant | 2.08 | 0.27 | | 0.56 | 0.31 | | 1.69 | 0.52 | |
| CNM pre-test score | **0.42** | **0.06** | **<.001** | **0.55** | **0.07** | **<.001** | **0.49** | **0.11** | **<.001** |
| Percent core activities | 0.03 | 0.19 | .549 | **0.20** | **0.24** | **<.001** | −0.02 | 0.40 | .865 |
| Percent of variance explained—total | 17.30 | | | 37.80 | | | 21.80 | | |
| Percent of variance explained—percent core activities | 0 | | | 3.40 | | | 0 | | |

[a]One item was removed from the MR construct for the biology domain because no students in the sample responded to this survey item.
*Note*. Dependent variable = epistemology score on each of the five SUMS subscales. Bold values indicate statistical significance for at least $p < .05$ (two-tailed).

models were associated with having engaged in a greater number of Connected Chemistry curricular activities.

*Analysis for Research Question 3: Are students' models pre-test scores a predictor of post-test content scores?*

To examine this question, we formulated regression models within each domain to examine whether students' pre-test model subscale scores were a predictor of their post-test content knowledge scores. In these models, students' pre-test content knowledge scores and the percentage of core activities engaged in by students were included as covariates. Each regression model is presented in Table 8.

Within the physics domain, the only significant predictor of students' post-test content domain scores was their pre-test content scores ($\beta = 0.65$, $p < .001$). The standardized regression coefficients associated with students' pre-test models scores on the five subscales were not significantly related to students' post-test content knowledge scores after controlling for students' pre-test content knowledge and the percentage of curricular activities used by students.

Within the biology domain, the standardized regression coefficient for the CNM pre-test subscale was positively and significantly related to students' post-test content knowledge ($\beta = 0.27$, $p < .05$) after controlling for students' pre-test content knowledge ($\beta = 0.37$, $p < .001$), the percentage of core activities ($\beta = 0.23$, $p < .05$), and the remaining four model pre-test subscale scores, none of which were significant. This suggests that holding all other variables in the model constant, a one unit increase in students' pre-test CNM scale was associated with a 0.27 standard deviation increase in students' post-test content knowledge.

For chemistry, two of the pre-test model subscale scores, namely, the MR ($\beta = 0.26$, $p < .05$) and the USM ($\beta = -0.31$, $p < .001$) subscales, were significantly associated with students' post-test knowledge, after controlling for students' pre-test content knowledge and percentage of core activities (both of which were statistically significant). Specifically, higher scores on the MR scale were associated with significantly higher post-test content knowledge scores ($\beta = 0.26$, $p < .05$); holding all other variables in the model constant, a one unit increase in students' pre-test MR subscale was associated with a 0.26 standard deviation increase in students' post-test content knowledge score. Conversely, holding all other variables in the model constant, a one unit increase in students' pre-test USM scale was associated with a 0.31 standard deviation decrease in students' post-test content knowledge.

## Discussion

This research was conducted with three research questions in mind. We have separated the discussion into three sections in which we reiterate the questions, summarize the findings, and then discuss these for each of the research questions, respectively.

Table 8. Regression models using epistemology pre-test subscales to predict post-test content knowledge

| Outcome variable: post-test content knowledge score | Physics ($N = 420$) | | | Chemistry ($N = 218$) | | | Biology[a] ($N = 98$) | | |
|---|---|---|---|---|---|---|---|---|---|
| | $\beta$ | *SE* | Significant | $\beta$ | *SE* | Significant | $\beta$ | *SE* | Significant |
| Constant | 8.099 | 2.05 | | –3.68 | 2.03 | | 4.04 | 4.91 | |
| Pre-test content knowledge score | **0.65** | **0.05** | **<.001** | **0.42** | **0.09** | **<.001** | **0.37** | **0.12** | **<.001** |
| % Core activities | 0.06 | 1.089 | .150 | **0.37** | **1.35** | **<.001** | **0.23** | **3.32** | **<.05** |
| ET pre-test score | 0.02 | 0.55 | .705 | 0.08 | 0.80 | .551 | –0.19 | 1.54 | .137 |
| ER pre-test score | –0.07 | 0.41 | .172 | –0.05 | 0.64 | .617 | –0.11 | 1.15 | .294 |
| MR pre-test score | –0.05 | 0.55 | .460 | **0.26** | **0.82** | **<.05** | 0.24 | 1.35 | .054 |
| USM pre-test score | 0.03 | 0.29 | .588 | **–0.31** | **0.45** | **<.001** | –0.12 | 0.91 | .247 |
| CNM pre-test score | 0.07 | 0.43 | .174 | 0.07 | 0.50 | .464 | **0.27** | **1.10** | **<.05** |
| Percent of variance explained | | 45.80 | | | 53.60 | | | 33.10 | |

[a]One item was removed from the MR construct for the biology domain because no students in the sample responded to this survey item.
*Note*. Bold values indicate statistical significance for at least $p < .05$ (two-tailed). The regression coefficients are standardized to have a mean of 0 and a standard deviation of 1.

*Domain-Specific Differences in Students' Understanding of Models*

First, we sought to test whether there were differences in students' pre-intervention understanding of models on each of the subscales measured across the three science domains. Regarding the measurement of students' understanding of models, some have claimed that assessing domain-specific differences here requires the use of domain-specific measurement tools (Hofer, 2000; Smith & Wenk, 2006). However, we feel that using a domain-general nature of models instrument that is administered in the context of three different science classes permits the rigor of using a single validated instrument *and* has the potential to capture and delineate the differences in students' understanding of models in different domains. This is consistent with what others have referred to as a bottom-up approach (Op't Eynde, De Corte, & Verschaffel, 2006). In the present study, the approach of using the same survey in three different science domains did yield differences across our three domains; thus, our conjecture about this appears to be, at least in part, correct.

Our findings show that there were significant differences in students' understanding of the nature of models prior to the implementation of the model-based curriculum in the three domains. Students who were about to engage in our model-based curricula in their physics class obtained model scores that represented higher levels of understanding, as measured by a 1–5 Likert scale, than did the biology students on three of the five subscales, namely, ET, MR, and USM. Additionally, students who were about to engage in their chemistry class obtained a model score that implied a more sophisticated understanding, as measured by a 1–5 Likert scale, than did the biology students for the USM subscale.

These data suggest that students' understanding of the nature of models differed across the three domains. Prior work in the broader area of epistemology (defined as nature of knowledge) has shown that there are differences across domains (Hofer, 2006a, 2006b; Muis, Bendixen, & Haerle, 2006), and thus, is compatible with our findings regarding the nature of models. Specifically, previous research has revealed differences in students' understanding of the nature of knowledge for different academic subjects such as science and psychology (Hofer, 2000), psychology and biology (Estes, Chandler, Horvath, & Backus, 2003), life sciences, analytic sciences, and the humanities (Royce & Mos, 1980), mathematics and social science (Paulsen & Wells, 1998), business, mathematics, and social science (Schommer-Aikins, Duell, & Barker, 2003), and history and mathematics (Buehl & Alexander, 2005). Our work contributes to the literature by substantiating that domain-specific differences in students' understanding of models exist across different disciplines of science.

Smith et al. (2000) claim that students develop different understandings for different domains because they encounter competing knowledge claims in these domains; Op't Eynde et al. (2006) concurs on this. In terms of our findings, it is possible that students' understanding of models in each domain is based on the models to which they have been exposed within these respective domains, and that these may have led to different understandings of models, which were elicited on the SUMS survey subscales (Treagust et al., 2002). Further research conducted in situ, i.e., within the context of each domain, using individual interviews may further delineate the differences in students' understanding of models in each of these domains. Research of this type which utilized one-on-one interviews has been successful at characterizing fine-grained differences in students' knowledge of models (Carey et al., 1989; Grosslight et al., 1991; Smith et al., 2000; Wenk & Smith, 2004), and thus this approach would likely bear fruit regarding disciplinary differences across science domains in students' understanding of models as well.

Given that we determined, as part of a secondary set of analyses, that students in the biology cohort were the youngest in the three groups, and that the chemistry students were the oldest in the cohort, we must be careful in interpreting our results to Research Question 1 as discipline-based differences. That is, a more complex explanation may be possible. For example, this pattern of results may also reflect differences in the students in the sample due to when students take specific science courses in high school. More specifically, in the USA, students typically take biology before chemistry or physics; this would mean that biology is the first science discipline students were exposed to in high school. In terms of our findings, it is possible that since biology was the first context in which their understanding of models was assessed, these students did have less sophisticated understandings of models (as measured by the 1–5 Likert scales), since they had taken fewer science courses. This interpretation would suggest that students are possibly fleshing out their understanding of models cumulatively as they encounter models in different science domains. Prior research has found that the amount of experience within a particular subject area appears to affect students' understanding about that domain (Schommer-Aikins et al., 2003); these findings, like ours, would suggest that students' understanding of models is affected by students' experiences in learning. These data are also consistent with the finding that the greatest gains in understanding of models were yielded in the chemistry group, which was the oldest cohort. Lastly, given the manner in which this study was conducted, i.e., via a survey instrument to assess students' understanding of models, it is impossible to determine whether the observed differences in model scores are due to exposure to science course, to age, or to some combination of both. Further research is necessary in a high school context in which biology is taken later in the high school curriculum than are physics and chemistry.



### *The Effects of Modeling on Students' Understanding of Models*

In our second research question we addressed whether engaging students in rich, authentic, model-based learning would influence their understanding of the nature of models. For biology and physics, there was no relationship found between the number of modeling activities the students engaged in and their post-test models scores. In chemistry, in which students were instructed about the nature and purpose of models in science, we found that the number of core activities was positively related to students' model scores at post-test; that is, doing more chemistry activities was associated with a more sophisticated understanding in all the subscales, as measured by a 1–5 Likert scale, with the exception of the ER subscale.

These results suggest that the explicit instruction of students' understanding of models in Connected Chemistry was effective at promoting students' understanding of models in this domain. Since explicit instruction was not provided in the physics and biology curricula, and there were no parallel gains yielded as a result of completing a greater number of modeling activities in physics and biology, our explanation of the role of explicit scaffolding in deepening students' understanding of models is consistent with our findings across the three domains.

Our findings here also appear to be compatible with those of prior research; that is, in cases in which changes in students' understanding of models resulted, the curricula were specifically designed with this goal. For example, Carey and her colleagues (Carey et al., 1989; Carey & Smith, 1993; Honda, 1994) directly taught students about the nature and purpose of models as part of various short-term curricular studies. Although changes in students' views about models were found, these were modest, at best (Smith et al., 2000). Schwarz and White (2005) showed that a curriculum, which was designed to teach about the nature of models, was successful at clarifying and broadening students' understanding of the nature and purpose of models. Lastly, Gobert and colleagues' curriculum (Gobert & Pallant, 2004; Gobert, Slotta et al., 2002; Gobert, Snyder, et al., 2002), which provided students with instruction about the nature of models in addition to engaging them in model-based inquiry tasks, yielded gains in students' open-response written questions about the nature and purpose of models.

In general, the pattern of results across these three domains yielded here provides some evidence that growth in students is possible, but that change is slow and needs to be explicitly scaffolded. Furthermore, given our results regarding the gains in chemistry, it is likely that a productive approach is to tightly align scaffolding with model-based tasks; a similar approach was also used in the curriculum design work by Gobert and her colleagues described above (Gobert & Pallant, 2004; Gobert, Slotta, et al., 2002; Gobert, Snyder, et al., 2002). We provide additional detail about how this might be accomplished in the 'Implications for Science Education' section.



### *The Effects of Understanding of Models on Content Learning*

For our third research question, we addressed whether any of the model subscales (as measured by the pre-test) were predictors of post-test content gains in each of the three content domains. This is based on the hypothesis that students' understanding of models might play a role in how they engage in learning science and, thus, might impact their understanding of the domain, as measured by the post-test in each domain. Similar claims have been made by others; for example, Sins et al. (2009) showed that those with a more sophisticated understanding of the nature of models were also those who engaged in deeper processing of science material.

To reiterate our findings to this question: in physics, none of the five subscales measuring students' understanding of the nature of models before the implementation had a significant relationship with the post-test scores in physics. In biology, a relationship was found between the subscale CNM and content knowledge, wherein those who had higher pre-test scores on the CNM also obtained higher content gains scores. In chemistry, a higher model pre-score was associated with a higher score on the chemistry content post-test for the MR scale; however, for the USM subscale, an inverse relationship was found; that is, those who scored lower on this model subscale tended to score higher on the content post-test. Each is addressed in turn.

With regard to physics, our findings suggest that differing levels of sophistication in students' understanding of models did not play a role in students' learning of content in this domain; i.e., those with a more sophisticated understanding of models did not necessarily learn more physics, as measured by our post-test, nor did those with a less sophisticated understanding of models necessarily learn less physics, as measured by our post-test.

With regard to biology, those who had a higher model score for the CNM at the pre-test, i.e., those who better understood that models change over time due to advances in understanding, also learned more biology, as measured by our post-test content items. The CNM subscale reflects the permanency or dynamic nature of models in science (Treagust et al., 2002). The items in this scale include 'A model may be changed if there are changes in data or beliefs'; 'A model may be changed if there are new findings'; and 'A model may be changed if new theories or evidence prove otherwise.' Based on our data, it is difficult to explain why significant findings on content learning relative to this subscale were observed in biology and not in the other two domains. In fact, these findings are particularly surprising since the curriculum for biology does not address the changing nature of models, nor does either of the two curricular packages. One possible explanation for this finding might be that the media attention on the human genome project provided students with some knowledge about how causal models of genetics have changed over time, and that this knowledge is being reflected in both the subscale for the CNM and students' content learning in the domain of biology. Specific questions to students addressing how the models in this domain have changed over time might provide some insight into whether or not media attention on the human genome project is a viable explanation for these findings.

With regard to chemistry, those who had higher scores on the subscale measuring MR also learned more chemistry, as measured by the post-test. This scale has six items, which address the reasons why multiple models may be used (different versions, different sides/parts of an object, different parts, how different parts are used, etc.). An example of a scaffold regarding multiple representations in the context of the chemistry curriculum is as follows:

> Scientists often develop different types of computer models to explore and understand the same complex system. Some of the models are more precise and some are more approximate. One reason that scientists develop a model with precise rules is that detailed models allow the study of the behavior of single objects in greater detail. At this point, you are going to switch from using a simplified model of gas particles in a container, to more precise models.

In this scaffold, we see that students are provided a rationale for why multiple models are used. Students, after reading this, are then presented with an explicit statement about the models they are then going to use to learn with. This scaffold is nicely aligned in the curriculum with its respective activity in which a new type of model is used to represent the phenomena under inquiry. One explanation for our findings here is that this scaffold, and others like it, provided conceptual leverage, especially for those who already had an understanding that models can have multiple representations in science. This, in turn, may have positively influenced their subsequent learning in the curriculum; evidence for this is the higher post-test content scores yielded for these students.

A second, curious finding was also observed regarding the Connected Chemistry curriculum regarding the relationship between students' understanding of models as measured at the pre-test and their content learning. Here, for the subscale USM, it was found that those who scored *lower* on this model subscale tended to score *higher* on the content post-test. Again, we looked into the nature of the questions in this subscale and the nature of the scaffolds that might have influenced students' learning in the curriculum. The SUMS questions that assessed this aspect of students' understanding of models are 'Models are used to make and test predictions about a scientific event' and 'Models are used to help formulate ideas and theories about scientific events.' In terms of scaffolding related to this aspect of models, an example is:

> You have explored a computer model to derive a set of relationships which govern the behaviors of gases in a container, such as the Ideal Gas Law. Your discovery of these important relationships shows how computer models can be used to make predictions and find quantitative relationships among variables. Imagine you are asked to help outline the requirements for a new computer model. This model will be used by weather forecasters to predict the temperature, pressure, and rainfall in different cities around the country. The behavior of the atmosphere depends a lot on the interactions among gas particles. The students are then asked: What are some of the important objects you would suggest including in the computer model? and What are some of the important properties of these objects that you would suggest including?

It is important to note as a caveat that one question from this subscale was dropped because its wording was problematic, thus our data on this scale is based on only two items (above), so we should be careful about over-interpreting our data. In the example of scaffolding above, students are told that a model can be used to 'make predictions and find quantitative relationships among variables'; in this task, we see a complex, albeit scientifically authentic task. Our data suggest that students who had a *less* sophisticated understanding of the uses of models *learned more chemistry*, as measured by the content post-test.

One possible explanation for this finding is that this scaffold provided a useful framework for students who had a naïve understanding of the uses of models, as measured by a 1–5 Likert scale, and that this helped them in this subsequent learning in the curriculum. In another study, a similar result was found. Specifically, in Wenk and Smith (2004), it was found that for students with lower model pre-test scores (Level 1.5/1.75 out of 3), science inquiry courses were quite effective in developing their understanding of models to a more advanced level. Additionally, the intervention was not as effective for students whose understandings of models at the onset were moderately sophisticated (as measured by a 2.0 out of 3 or 2.25/2.50 out of 3 on Wenk and Smith's scale). Although Wenk and Smith did not relate these model pre-test scores to students' post-test content scores, their findings are insightful when interpreting our data since it is possible that greater leverage was afforded by those who began the curriculum with a less sophisticated understanding of the USM, as measured by a 1–5 Likert scale. Further research, again using think-aloud protocols while the students work with our activities, might provide insight about how the scaffolds were used by students whose model pre-test scores differed in terms of their level of sophistication. For each type of subscale and its corresponding scaffolds, we would investigate how students of differing pre-test model scores made use of this scaffold to guide their learning in the curriculum. In the present research, this type of one-on-one testing was not conducted since the goal of the IERI program under which we were funded was that of scalability.



## Implications for Science Education

Since typical science instruction does not represent the real world of science and scientific practices, it is not surprising that students have naïve views of the nature of science, of scientific inquiry, and the nature of models (Carey et al., 1989; Chinn & Malhotra, 2002; Driver et al., 1996; Lederman, 1992). The research presented here, and that of others, has shown that students' understanding of models can be developed, but that this is difficult to achieve and needs to be directly scaffolded.

In the 'Introduction' section, we conceptualized the understanding of the nature of models as an integral subset to the understanding of the nature of science (Lederman, 2006). We base this on the important and inextricable coupling between models and the important role they play in scientific inquiry. Specifically, Hodson (1992) claims that learning about science, a critical component of scientific literacy, requires that students have an understanding of the nature of models and that they appreciate the role the models play in the accreditation and dissemination of scientific knowledge. From this, it follows that models and modeling need to play a central role in science education (Justi & Gilbert, 2002b). From our data, and that of others (cf. Schwarz & White, 2005; Smith et al., 2000; Treagust et al., 2002), it follows that students' understanding of models needs to be developed in order for students to learn successfully from and with models.

Based on prior literature, as well as our own research, we address some implications for science instruction that may serve to promote students' understanding of models. First, our research as well as that of others suggests that modeling practices should be explicitly taught. This is consistent with reform efforts, all of which emphasize modeling as an authentic scientific practice and its importance for science literacy (NRC, 1996). Modeling practices include creating, expressing, and testing models (Justi & Gilbert, 2002b). Furthermore, Justi and Gilbert (2002a) have clearly articulated a 'model of modeling' framework, which describes how modeling should be taught in classrooms so that learning is authentic for students. The components described by Justi and Gilbert are consistent with theories of model-based teaching and learning (cf. Gobert & Buckley, 2000). Specifically, in terms of direct instructional implications, Justi and Gilbert claim that students should be required to: (1) learn the use of models, that is, to explore scientific phenomena and conduct experiments, (2) learn to revise models with new evidence or feedback, and (3) learn how to construct models of scientific phenomena. It is believed that these instructional activities will foster students' modeling knowledge, including their understandings of the nature and purpose of models.

Schwarz, who also promotes the explicit teaching of models and modeling, describes a need for meta-modeling knowledge (Schwarz & Gwekwerere, 2006; Schwarz & White, 2005; Schwarz et al., 2007). Further, Schwarz and her colleagues prescribe that students need opportunities to engage and reflect on legitimate modeling experiences that are well aligned with content knowledge. She claims that the depth of students' understanding of the nature of models is likely to arise or emerge by having students deeply engage in modeling with a variety of inquiry tasks; others have also made similar claims and yielded data that supports this claim (Gobert & Discenna, 1997; Gobert & Pallant, 2004).

In another project currently underway (Gobert, Heffernan, Beck, & Koedinger, 2009; Gobert, Heffernan, Ruiz, & Kim, 2007; Sao Pedro, Gobert, Beck, & Heffernan, 2009), students are engaged in scientific inquiry using models, i.e., microworlds, in order to make predictions, design and test experiments, interpret data, and compare data with their predictions. In this project, the microworlds serve as tools that provide critical, perceptual and conceptual affordances for students to hone both their content knowledge and their inquiry skills (Gobert, 2005; Sao Pedro et al., 2009). It is our belief that scaffolding authentic modeling activities will promote students' understanding of models as well as of content knowledge. We seek to address this important question as our work on this project unfolds, thereby contributing to this important area of science education.



## Acknowledgments

This research was funded by the National Science Foundation under the Interagency Education Research Initiative (IERI no. 0115699) awarded to The Concord Consortium and to Northwestern University. Any ideas or opinions expressed are those of the authors and do not necessarily reflect the views of the sponsoring agency. This work was also supported by two additional grants, namely, one from the National Science Foundation (NSF-DRL no. 0733286) awarded to Janice D. Gobert, Neil Heffernan, Carolina Ruiz, and Ryan Baker, and one from the U.S. Department of Education (R305A090170) awarded to Janice D. Gobert, Neil Heffernan, Ken Koedinger, and Joseph Beck. All opinions expressed do not necessarily reflect the views of the agency. The authors wish to thank Christina Schwarz, Barbara Hofer, and David Hammer on their helpful comments on the final version of this manuscript.

## Notes

1. The Connected Chemistry curriculum has gone through several iterations and ensuing versions, starting with Wilensky's (1999b) and GasLab, Stieff and Wilensky's (2003) version. The version used in the work reported on herein is known as CC1 (Levy et al., 2004; Levy & Wilensky, 2009).
2. Note that the Treagust et al.'s (2002) original SUMS survey comprised 27 items. One item from the USM subscale was removed because the wording was deemed problematic.

## References

Abd-El-Khalick, F., & Lederman, N. G. (2000). Improving science teachers' conceptions of the nature of science: A critical review of the literature. *International Journal of Science Education, 22*(7), 665–701.

Bell, P. (1998). *Designing for students' science instruction using argumentation to classroom debate.* Unpublished doctoral dissertation, University of California, Berkeley.

Bisard, W. J., Aron, R. H., Francek, M., & Nelson, B. D. (1994). Assessing selected physical science and earth science misconceptions of middle school through university pre-service teachers. *Journal of College Science Teaching, 24*(1), 38–42.

Buehl, M., & Alexander, P. (2005). Students' domain-specific epistemological belief profiles. *American Educational Research Journal, 42*(4), 697–726.

Burbules, N. C., & Linn, M. C. (1991). Science education and philosophy of science: Congruence or contradiction? *International Journal of Science Education, 22*(8), 798–817.

Carey, S., Evans, R., Honda, M., Jay, E., & Unger, C. (1989). An experiment is when you try it and see if it works: A study of grade 7 students' understanding of the construction of scientific knowledge. *International Journal of Science Education, 11*, 514–529.

Carey, S., & Smith, C. (1993). On understanding the nature of scientific knowledge. *Educational Psychologist, 28*, 235–251.

Carey, S., & Smith, C. (1995). On understanding the nature of scientific knowledge. In D. Perkins, J. Schwartz, M. Maxwell West, & M. Stone Wiske (Eds.), *Software goes to school* (pp. 39–55). Oxford: Oxford University Press.

Champagne, A., Gunstone, R., & Klopfer, L. (1985). Effecting changes in cognitive structures among physics students. In L. West & A. Pines (Eds.), *Cognitive structure and conceptual change* (pp. 61–90). New York: Academic Press.

Chen, Z., & Klahr, D. (1999). All other things being equal: Acquisition and transfer of the control of variables strategy. *Child Development, 70*(5), 1098–1120.

Chinn, C. A., & Malhotra, B. A. (2002). Epistemologically authentic reasoning in schools: A theoretical framework for evaluating inquiry tasks. *Science Education, 86*, 175–218.

Chittleborough, G., Treagust, D., Mamiala, M., & Mocerino, M. (2005). Students' perceptions of the role of models in the process of science and in the process of learning. *International Journal of Science Education, 23*(2), 195–212.

Clement, J. (1993). Model construction and criticism cycles in expert reasoning. In *The proceedings of the 15th annual conference of the Cognitive Science Society* (pp. 336–341). Hillsdale, NJ: Lawrence Erlbaum.

Clement, J., Brown, B., & Zietsman, A. (1989). Not all preconceptions are misconceptions: Finding 'anchoring conceptions' for grounding instruction on students' intuitions. *International Journal of Science Education, 11*, 554–565.

Crawford, B. A., & Cullin, M. J. (2004). Supporting prospective teachers' conceptions of modeling in science. *International Journal of Science Education, 26*(11), 1379–1401.

Driver, R., Guesne, E., & Tiberghien, A. (1985). *Children's ideas in science.* Buckingham: Open University Press.

Driver, R., Leach, J., Millar, R., & Scott, P. (1996). *Young people's images of science.* Buckingham: Open University Press.

Driver, R., Squires, A., Rushworth, P., & Wood-Robinson, V. (1994). *Making sense of secondary science.* New York: Routledge.

Estes, D., Chandler, M., Horvath, K. J., & Backus, D. W. (2003). American and British college students' epistemological beliefs about research on psychological and biological development. *Applied Developmental Psychology, 23*, 625–642.

Giere, R. N. (1990). *Explaining science.* Chicago: University of Chicago Press.

Gilbert, J. K. (Ed.). (1993). *Models and modelling in science education.* Hatfield: Association for Science Education.

Gobert, J. (2000). A typology of models for plate tectonics: Inferential power and barriers to understanding. *International Journal of Science Education, 22*(9), 937–977.

Gobert, J. (2005). Leveraging technology and cognitive theory on visualization to promote student's science learning and literacy. In J. Gilbert (Ed.), *Visualization in science education* (pp. 73–90). Dordrecht: Springer-Verlag.

Gobert, J., & Buckley, B. (2000). Special issue editorial: Introduction to model-based teaching and learning. *International Journal of Science Education, 22*(9), 891–894.

Gobert, J., Buckley, B., & Clarke, J. (2004, April). *Scaffolding model-based reasoning: Representations, information-processing, and cognitive affordances.* Paper presented at the annual meeting of the American Educational Research Association, San Diego, CA.

Gobert, J., & Clement, J. (1999). Effects of student-generated diagrams versus student-generated summaries on conceptual understanding of causal and dynamic knowledge in plate tectonics. *Journal of Research in Science Teaching, 36*(1), 39–53.

Gobert, J., & Discenna, J. (1997). *The relationship between students' epistemologies and model-based reasoning.* Kalamazoo: Department of Science Studies, Western Michigan University. (ERIC Document Reproduction Service No. ED409164.)

Gobert, J., Heffernan, N., Beck, J., & Koedinger, K. (2009). ASSISTments Meets Science Learning. Proposal funded by the US Department of Education (R305A090170).

Gobert, J., Heffernan, N., Ruiz, C., & Kim, R. (2007). AMI: ASSISTments Meets Inquiry. Proposal funded by the National Science Foundation (NSF-DRL# 0733286).

Gobert, J., & Horwitz, P. (2002). Do modeling tools help students learn science? *In @ Concord, 6*(1), 19.

Gobert, J. D., & Pallant, A. (2004). Fostering students' epistemologies of models via authentic model-based tasks. *Journal of Science Education and Technology, 13*(1), 7–22.

Gobert, J., Slotta, J., & Pallant, A. (2002, April). *Inquiry learning through students' east-west coast collaboration.* Paper presented at the National Association for Research in Science Teaching, New Orleans, LA.

Gobert, J., Snyder, J., & Houghton, C. (2002, April). *The influence of students' understanding of models on model-based reasoning.* Paper presented at the annual meeting of the American Educational Research Association, New Orleans, LA.

Grosslight, L., Unger, C., Jay, E., & Smith, C. (1991). Understanding models and their use in science: Conceptions of middle and high school students and experts. *Journal of Research in Science Teaching, 28*(9), 799–822.

Hammer, D. (1994). Epistemological beliefs in introductory physics. *Cognition and Instruction, 12*(2), 151–183.

Hammer, D. (1995). Epistemological considerations in teaching introductory physics. *Science Education, 79*(4), 393–413.

Hammer, D., & Elby, A. (2002). On the form of a personal epistemology. In B. Hofer & P. Pintrich (Eds.), *Personal epistemology: The psychology of beliefs about knowledge and knowing* (pp. 169–190). Mahwah, NJ: Lawrence Erlbaum.

Hennessey, M. G. (1995, September). *Students' epistemological stance: Nature of learning and nature of science.* Paper presented at the Cognitive Studies and Educational Practice Meetings of the McDonnell Foundation, Nashville, TN.

Hennessey, M. G., & Beeth, M. (1993, April). *Students' reflective thoughts about science content: A relationship to conceptual change learning.* Paper presented at the meeting of the American Educational Research Association, Atlanta, GA.

Hesse, M. B. (1963). *Models and analogies in science.* London: Sheed & Ward.

Hestenes, D. (1987). Toward a modeling theory of physics instruction. *American Journal of Physics, 55*(5), 440–454.

Hodson, D. (1992). In search of a meaningful relationship: An exploration of some issues relating to integration in science and science education. *International Journal of Science Education, 14,* 541–562.

Hofer, B. (2000). Dimensionality and disciplinary differences in personal epistemology. *Contemporary Educational Psychology, 25*, 378–405.

Hofer, B. (2006a, May). *Understanding students' epistemic beliefs in math and science: An overview of constructs, measures, and research.* Paper presented at the BRAIN Conference (Broadening Research at International Network: Developing a Cross-Domain Research Framework for Science and Math Education), National Taiwan Normal University, Taipei.

Hofer, B. (2006b). Domain specificity of personal epistemology: Resolved questions, persistent issues, new models. *International Journal of Educational Research, 45*, 85–95.

Honda, M. (1994). *Linguistic inquiry in the science classroom: 'It is science, but it's not like a science problem in a book'* (Working Papers in Linguistics). Cambridge, MA: MIT Press.

Justi, R. S., & Gilbert, J. K. (2002a). Modelling, teachers' views on the nature of modeling, and implications for the education of modelers. *International Journal of Science Education, 24*(4), 369–388.

Justi, R. S., & Gilbert, J. K. (2002b). Science teachers' knowledge about and attitudes towards the use of models and modeling in learning science. *International Journal of Science Education, 24*(12), 1273–1292.

Larkin, J., & Simon, H. (1987). Why a diagram is (sometimes) worth ten thousand words. *Cognitive Science, 11*, 65–100.

Lederman, N. G. (1992). Students' and teachers' conceptions of the nature of science: A review of the research. *Journal of Research in Science Teaching, 29*(4), 331–359.

Lederman, N. G. (2006). Syntax of nature of science within inquiry and science instruction. In L. B. Flick & N. G. Lederman (Eds.), *Scientific inquiry and nature of science: Implications for teaching, learning, and teacher education* (pp. 301–317). Dordrecht: Springer.

Levy, S. T., Kim, H., & Wilensky, U. (2004, April). *Connected Chemistry—A study of secondary students using agent-based models to learn chemistry.* Paper presented at the annual meeting of the American Educational Research Association, San Diego, CA.

Levy, S. T., & Wilensky, U. (2009). Crossing levels and representations: The Connected Chemistry (CC1) curriculum. *Journal of Science Education and Technology, 18*(3), 224–242.

Linn, M. C., & Muilenberg, L. (1996). Creating lifelong science learners: What models form a firm foundation? *Educational Researcher, 25*(5), 18–24.

Linn, M., Songer, N. B., Lewis, E. L., & Stern, J. (1991). Using technology to teach thermodynamics: Achieving integrated understanding. In D. L. Ferguson (Ed.), *Advanced technologies in the teaching of mathematics and science* (pp. 5–60). Berlin: Springer-Verlag.

Lowe, R. (1989). *Scientific diagrams: How well can students read them? What research says to the science and mathematics teacher* (Vol. 3). Perth: Key Centre for School Science and Mathematics, Curtin University of Technology.

Lowe, R. (1993). Constructing a mental representation from an abstract technical diagram. *Learning & Instruction, 3*, 157–179.

Muis, K., Bendixen, L., & Haerle, F. (2006). Domain-generality and domain-specificity in personal epistemology research: Philosophical and empirical reflections in the development of a theoretical framework. *Educational Psychology Review, 18*, 3–54.

NRC (National Research Council). (1996). *National science education standards: 1996.* Washington, DC: National Academy Press.

Op't Eynde, P., De Corte, E., & Verschaffel, L. (2006). Epistemic dimensions of students' mathematics-related belief systems. *International Journal of Educational Research, 45*(1–2), 57–70.

Paulsen, M. B., & Wells, C. T. (1998). Domain differences in the epistemological beliefs of college students. *Research in Higher Education, 39*(4), 365–384.

Perkins, D., Jay, E., & Tishman, S. (1993). Teaching thinking: From ontology to education. *Educational Psychologist, 28*(1), 67–85.

Perry, W. G., Jr. (1970). *Forms of intellectual and ethical development in the college years: A scheme.* New York: Holt, Rinehart, and Winston.

Posner, G. J., Strike, K. A., Hewson, P. W., & Gertzog, W. A. (1982). Accommodation of a scientific conception: Toward a theory of conceptual change. *Science Education, 66*(2), 211–227.

Raghavan, K., & Glaser, R. (1995). Model-based analysis and reasoning in science: The MARS curriculum. *Science Education, 79*(1), 37–61.

Royce, J. R., & Mos, L. P. (1980). *Manual: Psycho-epistemological profile.* Alberta: Center for Advanced Study in Theoretical Psychology, University of Alberta.

Sao Pedro, M., Gobert, J., Beck, J., & Heffernan, N. (2009, July). *Can an intelligent tutor teach the control of variables strategy for scientific inquiry*? Paper presented at the Cognitive Science Society Conference, Amsterdam.

Schommer-Aikins, M., Duell, O., & Barker, S. (2003). Epistemological beliefs across domains using Biglan's classification of academic disciplines. *Research in Higher Education, 44*(3), 347–366.

Schwarz, C. (2002, April). *Is there a connection? The role of meta-modeling knowledge in learning with models*. Paper presented at the International Conference of the Learning Sciences, Seattle, WA.

Schwarz, C. V., & Gwekwerere, Y. N. (2006). Using a guided inquiry and modeling instructional framework (EIMA) to support pre-service K-8 science teaching. *Science Education, 91*(1), 158–186.

Schwarz, C. V., Meyer, K., & Sharma, A. (2007). Technology, pedagogy, and epistemology: Opportunities and challenges of using computer modeling and simulation tools in elementary science methods. *Journal of Science Teacher Education, 18*(2), 243–269.

Schwarz, C., & White, B. (1999, March). *What do seventh grade students understand about scientific modeling from a model-oriented physics curriculum?* Paper presented at the National Association for Research in Science Teaching, Boston.

Schwarz, C., & White, B. (2005). Meta-modeling knowledge: Developing students' understanding of scientific modeling. *Cognition and Instruction, 23*(2), 165–205.

Sengupta, P., & Wilensky, U. (2009). Learning electricity with NIELS: Thinking with electrons and thinking in levels. *International Journal of Computers for Mathematical Learning, 14*(1), 21–50.

Sins, P. H. M., Savelsbergh, E. R., van Joolingen, W., & van Hout-Wolters, B. (2009). The relation between students' epistemological understanding of computer models and their cognitive processing in a modeling task. *International Journal of Science Education, 31*(9), 1205–1229.

Smith, C., Maclin, D., Houghton, C., & Hennessy, G. (2000). Sixth-grade students' epistemologies of science: The impact of school science experiences on epistemological development. *Cognition and Instruction, 18*(3), 349–422.

Smith, C., & Wenk, L. (2003, March). *The relation among three aspects of college freshman's epistemology of science.* Paper presented at the National Association for Research in Science Teaching, Philadelphia.

Smith, C., & Wenk, L. (2006). Relations among three aspects of first-year college students' epistemologies of science. *Journal of Research in Science Teaching, 43*(8), 747–785.

Snir, J., Smith, C., & Grosslight, L. (1988). *Not the whole truth: An essay on building a conceptually enhanced computer simulation for science teaching* (Tech. Rep. No. TR 88-18). Cambridge, MA: Educational Technology Center, Harvard Graduate School of Education.

Songer, N. B., & Linn, M. C. (1991). How do students' views of science influence knowledge integration? *Journal of Research in Science Teaching, 28*(9), 761–784.

Stieff, M., & Wilensky, U. (2002). ChemLogo: An emergent modeling environment for teaching and learning chemistry. In P. Bell, R. Stevens, & T. Satwicz (Eds.), *Keeping learning complex: The proceedings of the fifth international conference of the learning sciences (ICLS)* (pp. 451–458). Mahwah, NJ: Lawrence Erlbaum.

Stieff, M., & Wilensky, U. (2003). Connected Chemistry—Incorporating interactive simulations into the chemistry classroom. *Journal of Science Education & Technology, 12*(3), 285–302.

Strike, K., & Posner, G. (1985). A conceptual change view of learning and understanding. In L. H. T. West & A. L. Pines (Eds.), *Cognitive structure and conceptual change* (pp. 189–210). New York: Academic Press.

Treagust, D., Chittleborough, G., & Mamiala, T. (2001, April). *Learning introductory organic chemistry: Secondary students' understanding of the role of models and the development of scientific ideas.* Paper presented at the American Educational Research Association, Seattle, WA.

Treagust, D., Chittleborough, G., & Mamiala, T. (2002). Students' understanding of the role of scientific models in learning science. *International Journal of Science Education, 24*(4), 357–368.

van Driel, J. H., & Verloop, N. (1999). Teachers' knowledge of models and modeling in science. *International Journal for Science Education, 21*(11), 1141–1153.

Wenk, L., & Smith, C. (2004, April). *The impact of first-year college science courses on epistemological thinking: A comparative study.* Paper presented at the National Association of Science Teachers, Vancouver.

White, B., & Frederiksen, J. (1990). Causal model progressions as a foundation for intelligent learning environments. *Artificial Intelligence, 24*, 99–157.

Wilensky, U. (1999a). *NetLogo.* Evanston, IL: Center for Connected Learning and Computer-Based Modeling, Northwestern University. Retrieved March 30, 2010, from http://ccl.northwestern.edu/netlogo

Wilensky, U. (1999b). GasLab: An extensible modeling toolkit for exploring micro-and-macro-views of gases. In N. Roberts, W. Feurzeig, & B. Hunter (Eds.), *Computer modeling and simulation in science education* (pp. 151–178). Berlin: Springer-Verlag.

Wilensky, U. (2001). *Modeling nature's emergent patterns with multi-agent languages.* Paper presented at the Eurologo 2001 Conference, Linz, Austria. Retrieved March 30, 2010, from http://ccl.northwestern.edu/papers/MEE/

Wilensky, U., & Reisman, K. (2006). Thinking like a wolf, a sheep or a firefly: Learning biology through constructing and testing computational theories. *Cognition and Instruction, 24*(2), 171–209.