How do I get there... straight, oscillate or inch?
High-school students' exploration patterns of Connected Chemistry models

Sharona T. Levy & **Uri Wilensky**

stlevy@construct.haifa.ac.il
University of Haifa, Israel

uri@northwestern.edu
Center for Connected Learning
and Computer-Based Modeling
Northwestern University,
Evanston, IL

****This is a draft. Please do not circulate or quote without permission of the authors.****

Paper to be presented at the 2007 annual meeting of the
American Educational Research Association, Chicago, USA.

ABSTRACT

Connected Chemistry (Levy, Novak & Wilensky, 2005; 2006) is a curriculum package, designed for high-school students, as part of the Modeling Across the Curriculum (MAC) project (Gobert, Horwitz et al, 2003), in which students explore computer models of scientific phenomena. The Connected Chemistry curriculum focuses on topics in chemistry and employs multi-agent NetLogo models (Wilensky, 1999a) embedded in a supporting script to enable students in self-directed inquiry: manipulating and observing interactions between objects at the molecular level in order to gain insight into emergent patterns and macroscopic phenomena. Chapter 1 of connected chemistry focuses on the topic of gases: the microscopic particle behaviors in terms of Kinetic Molecular Theory; the macroscopic patterns, described by the Gas Laws; the relationship between these two description levels; and, the dynamics of change in the system.

The multi-agent NetLogo models are embedded within a Pedagogica™ (Horwitz & Burke, 2002) script. This integration of NetLogo and Pedagogica targets several goals, serving curriculum designers, students, as well as researchers. In authoring scripted activities, curriculum designers can combine the best of two worlds: the powerful modeling and exploration tools in NetLogo, and the interactive structured dialogue between the student and the embedding script via Pedagogica. In terms of assessment and research, Pedagogica enables the logging of students' actions within this learning environment. These include students' actions with the models, and their textual responses while predicting, observing, analyzing and reflecting upon the issues raised in the curriculum. This logging provides a rich set of data, which is mined, parsed and converted into research-able behaviors relating into the students' model manipulation, concurrent and overall learning.

In previous work, we have investigated students' patterns of exploration of the NetLogo models (Levy & Wilensky, 2005, 2006a, 2006b). We identified and described three prominent patterns: "straight to the point", "oscillation between extremes" and "inching through space". Herein, we show results of further analysis of these patterns on a larger data-set, in which the patterns were re-conceptualized by converting them into dimensions that characterize the students' explorations. We present a case study of one student interacting the models and the embedding scripts. We use this case to demonstrate the affordances of logging her activities for gaining a deeper understanding of the process by which she learns through interaction with the models. We demonstrate the particular affordances of the Connected Chemistry curriculum regarding micro-macro relations, and evolving mathematical modeling abilities through data on the pre- to post-test scores. We elaborate on the affordances of the integration between NetLogo and Pedagogica in authoring and designing activities.

INTRODUCTION

Learning about gases through model exploration

A body of science education literature points to student's misunderstandings of the gaseous phase of matter (Lin & Cheng, 2000; Maz & Perez, 1987). Some of these misunderstandings can be related to what Wilensky and Resnick call "levels confusion" (1999), where the properties of the macro-level are incorrectly ascribed to the micro-level (in the particular case of chemistry). The macroscopic properties of gases are easier to experience and perceive, such as when a kettle boils or a coke bottle produces a hiss when it's opened. However, the microscopic particles that are moving, colliding and bouncing off the walls are invisible. The literature reports a variety of student's non-standard conceptions about gases such as ordered packing of molecules and weightlessness of the gas. Lin and Cheng (2000) describe high-school students' difficulties in understanding Kinetic Molecular Theory as it applies to gases: molecules are pushed down, molecules stay away from heat and molecules expand when they are heated. All three can be related to our macroscopic daily experiences: the force of gravity pulling objects towards the earth, boiling water rising out of a pot and expansion of matter upon heating. Mas and Perez (1987) have found that high-school students regard gases as weight-less, reasoning from their observations that gases rise, and inferring that they therefore cannot have weight. Similar problems have been reported in a variety of scientific domains, such as genetics (Marbach-Ad & Stavy, 2000) and basic electricity concepts (Frederiksen, White & Gutwill, 1999).

The learning research community has recognized a disconnect between conceptual and algorithmic understandings of Chemistry (e.g., Kozma et al, 1996; Niaz & Robinson, 1992; Stieff & Wilensky, 2003). For example, Berg and Treagust (1993) point to the minimal use of qualitative relationships regarding teaching the gas laws both in a variety of textbooks they analyzed and in teaching approaches in schools. Students may be capable of solving problems that involve the procedures commonly taught in science classes. However, they do not necessarily do as well when approaching a similar problem that requires more qualitative, or conceptual reasoning.

A fruitful way of approaching the problem of bridging the conceptual and symbolic forms of representing chemical phenomena is to use computer models that employ multiple representations and that have affordances that enable connecting the representations (see 4M:Chem, Kozma et al, 1996). Frederiksen, White & Gutwill (1999) have employed a variety of conceptual models to design computer simulations to help students connect the different levels that can be used to describe basic electricity: a particle model, a circuit model and an algebraic model.

Wilensky and colleagues (Wilensky, 1999b; 2003; Wilensky, Hazzard & Froemke, 1999) have shown that NetLogo models can be powerful avenues for learning about gases and, more generally, about statistical mechanics. In their studies, students used the GasLab (Wilensky, 2000) package. Students were involved at three levels: exploring existing GasLab models, modifying those models, and constructing new such models.

The work reported here builds upon this previous work, but differs in that all the students are involved only at the exploratory level and that their explorations are guided and constrained by a script. The script is designed to guide but also to enable freedom and exploratory flexibility. However, the Connected Chemistry models do enable students to view (and modify) the underlying rules that generate the model behaviors. The affordance for students to connect the observed phenomena with the mechanism or rules underlying the model enables students to view the model as modifiable by them and not a prepared "movie" selected by the designers.

Connected Chemistry activities

Chemistry is a natural domain for an agent-based approach, as all chemical phenomena emerge from local interactions among a multitude of interacting individual molecules. In the Modeling Across the Curriculum project, we enable students' exploration of computer models that are embedded in a supporting script. The Connected Chemistry learning environment (Levy, Novak & Wilensky, 2005, 2006; Stieff & Wilensky, 2003) is one such model-based curricular unit. Connected Chemistry employs multi-agent NetLogo (models to empower the students' manipulation and observation of chemical "entities" at the molecular level as well as the resulting aggregate patterns. In this project, the models are embedded within a Pedagogica ^{TM}script (Horwitz & Burke, 2002) that structures the interaction of the students with the models, guide the model exploration as well as asking students questions about their exploration and findings. The first Connected Chemistry unit is on the topic of gases: Gas laws, and Kinetic Molecular Theory. The models used in the current project are a modified version of those originally created for the GasLab curriculum (Wilensky, 1999b).

The first set of activities in the Connected Chemistry curriculum is on the topic of gases: Gas laws, and Kinetic Molecular Theory (KMT). Kinetic Molecular Theory describes the behavior of individual particles (e.g., particles move in straight lines, they elastically collide with each other and with the walls). Gas laws describe the relationships among properties of the system of particles as a whole, when it is in equilibrium (e.g., Boyle's Law: the relationship between the volume of a box and the pressure inside, when temperature and the number of particles are constant). In addition to the traditional chemistry content, our curriculum also targets several important chemistry-related ideas: (a) Modeling: how a model is constructed, its assumptions, affordances and limitations, its relation with the target real-world phenomenon; (b) Thinking "from the molecule up" by focusing on micro-to-macro descriptions, transitions and connections; (c) Focus on processes of change in the system, such as perturbation and equilibration; (d) Mathematical modeling, deriving equations from data obtained through the students' NetLogo model explorations.

More generally, the chemistry topics are set within a wider perspective of complex systems. The domain of "complex systems" has evolved rapidly in the past 15 years, developing novel ideas and tools, and new ways of comprehending old phenomena, such as weather systems. Complex systems are made up of many elements (sometimes described as "agents", in our case, molecules), which interact among themselves and with their environment. The interactions of numerous elements result in a higher-order or collective behavior. Although such systems are not regulated through central control, they self-organize in coherent global patterns (Holland, 1995; Kauffman, 1995; Resnick & Wilensky, 1993). These patterns are often counter-intuitive and surprising.

The Connected Chemistry unit consists of a sequence of seven activities. The sequence of activities is as follows:

(1) **Modeling a Tire**: A rule-by-rule construction of the gas model, leading up to a focus on the Kinetic Molecular Theory (KMT) assumptions.

(2) **Changing pressure**: Introduces the concept of pressure, elaborating on processes of change, delays between perturbing the system until the system reacts and then re-equilibrates, relations between the randomness of the gas particles' motion and the stability of pressure.

(3) **Experimenting with particles**: New tools are offered in this activity – the use of several NetLogo commands to change the particles' properties, enhance and change the visual representations; propagating global effects from a local change; The students design and conduct an experiment of their choice, determining their course of action and using NetLogo commands to conduct their exploration.

(4) **Number and pressure**: The relationship between the number of particles in a fixed container and the pressure is explored, both qualitatively and quantitatively – deriving the equation that relates the two variables.

(5) **Temperature and pressure**: The concept of energy is elaborated upon via the changes to the gas temperature; the qualitative and quantitative relationship between temperature and pressure is investigated.

(6) **Volume and pressure**: The concept of pressure is further explored in this activity, as it relates to the area of the container's surface; the qualitative and quantitative relationship between the two variables is probed and summarized.

(7) **Ideal gas law**: Through both open investigation of a more complex gas model, and a guided mathematical derivation, the unit culminates in the Ideal Gas Law.

We portray the curriculum by presenting sample screenshots demonstrating two main goals: (1) Thinking "from the molecule up" by focusing on micro-to-macro descriptions, transitions and connections; (2) Mathematical modeling, deriving equations from data obtained through the students' NetLogo model explorations.

Thinking "from the molecule up"

Within an agent-based perspective of systems in general, as well as in chemistry, at least two distinct levels of description are necessary to make sense of phenomena. There is a two-way interaction between these two levels. Moving from the molecules upwards, a causal explanation of observed phenomena is made through their molecular descriptions. Molecules can be described via their behaviors and interactions. Through modeling of the multiple interacting molecules, these local behaviors emerge into coherent patterns of system-wide phenomena. However, the impetus for exploration is in the observed phenomena, when real-world events and situations beg for explanation. Thus, the curriculum moves both ways: from the phenomenon to its particles' behaviors, and vice versa.

To depict the curriculum as a whole, we have coded each question addressed to the students as one of four, with respect to the following:

(1) *Microscopic level*: Only molecular rules and behaviors are referenced. Examples: deriving gas particle rules, such as "moving in straight lines, unless they collide with something"; Describing changes in particle behaviors when they are allowed to collide with each other.

(2) *Macroscopic level*: Only group-wide patterns and variables are addressed. Examples: when learning about the gas laws, relating any two macroscopic variables, qualitatively or quantitatively; relating symbolic representations and qualitative changes in the system.

(3) *Micro-to-Macro*: Involves reasoning "up" from the molecular behaviors to the system's patterns and variables. Examples: relating changes to all the particles when the particles are "instructed" to bounce off the container's walls, or including gravity in the model; Connecting the changes in particle behavior due to gravity, to macroscopic phenomena, such as differences of pressure at different altitudes.

(4) *Macro-to-Micro*: Requires reasoning "down" from change at the level of the system to local molecular behaviors. Examples: relating pumping up a bicycle tire to the air particles' behavior; relating changes in the particles' average speed to energy transfer through their collisions.

Figures 1 and 2 (below) describes the seven activities in the curriculum with respect to the Levels dimension.

From these figures, we conclude:

(1) The curriculum shifts from a more microscopic perspective in the earlier activities to a more macroscopic perspective in the later activities.

(2) Transitions among levels (micro-to-macro, macro-to-micro), which target agent-based causal reasoning in the system, are frequent and occur throughout the activities. The third activity, which ends the first more qualitative section is dominant in highlighting these transitions. Among the two, micro-to-macro questions are more dominant, accenting an agent-based approach.

(3) The curriculum is made up of a high proportion of macro-level questions. This results from a focus in the second part on the students' mathematical modeling,

constructing the gas laws in a guided framework.

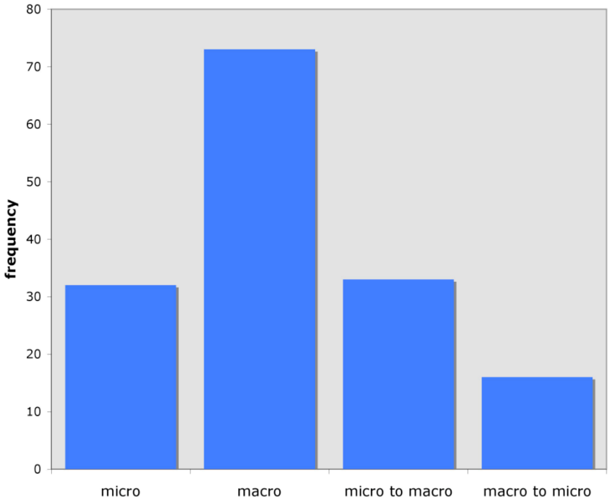


Figure 1: Connected Chemistry text-based questions – frequency of the different questions in the curriculum along the levels dimension.

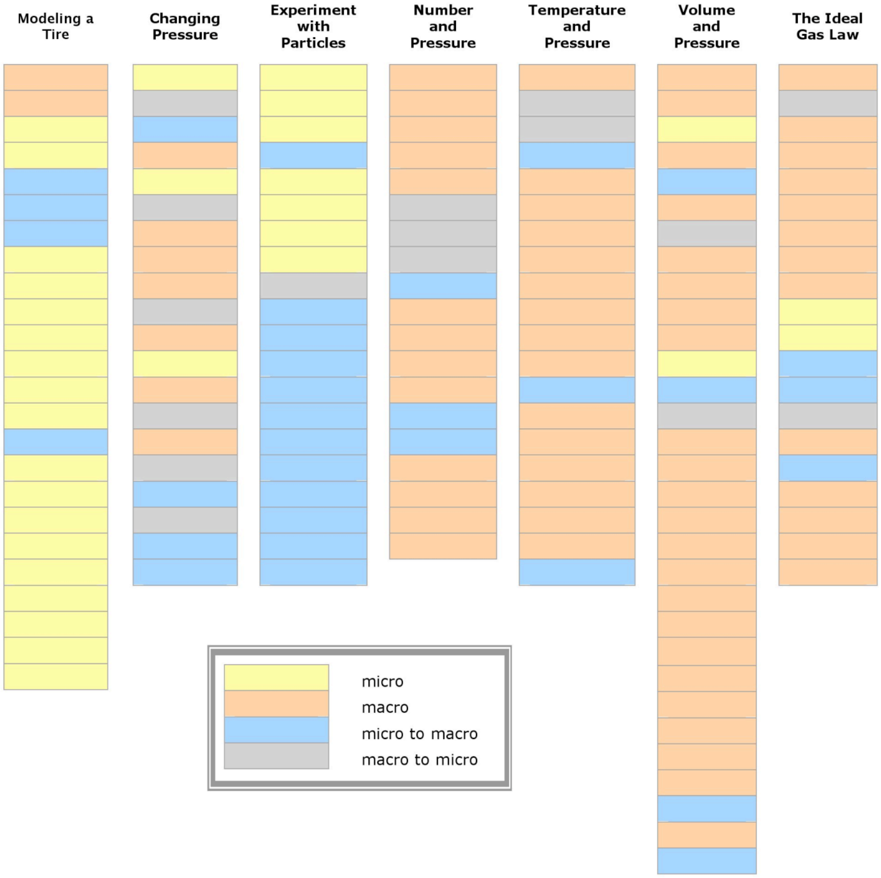
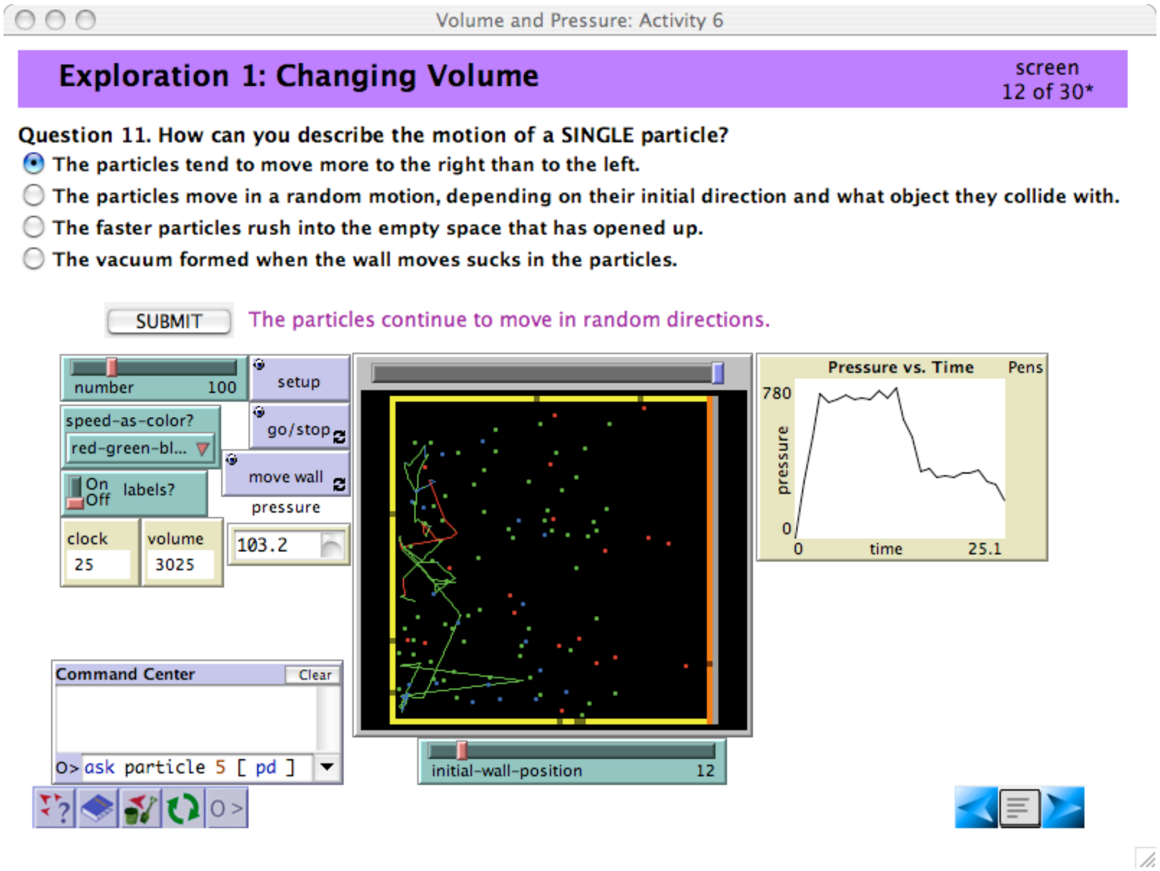


Figure 2: The “Levels” dimension in the Connected Chemistry curriculum. Each column is an activity. Each bar is a question addressed to the student. The bar is colored and illustrated via one of the four categories for levels: micro, macro, micro-to-macro, macro-to-micro.

We demonstrate the Levels dimension, “thinking from the molecule up”. Several explorations throughout the sequence call attention to microscopic particle behaviors and their relation to the system-wide variables. In the following activity (Figure 3), a model of a box containing gas affords changing its volume and observing several outcomes and behaviors. In this particular activity, the student is encouraged to follow a single particle by using the command line to mark its path. When the student increases the volume, one can note both micro- and macro- behaviors that are distinct: the particle continues its straight-line motion, which changes direction and speed only upon collisions; however, at a macro-level, the group of particles seems to be moving from left to right. The embedding script encourages noting the micro-behaviors.



Volume and Pressure: Activity 6

Exploration 1: Changing Volume

screen 12 of 30*

Question 11. How can you describe the motion of a SINGLE particle?

- ● The particles tend to move more to the right than to the left.
- ○ The particles move in a random motion, depending on their initial direction and what object they collide with.
- ○ The faster particles rush into the empty space that has opened up.
- ○ The vacuum formed when the wall moves sucks in the particles.

SUBMIT The particles continue to move in random directions.

number 100

speed-as-color? red-green-bl...

clock 25

volume 3025

pressure 103.2

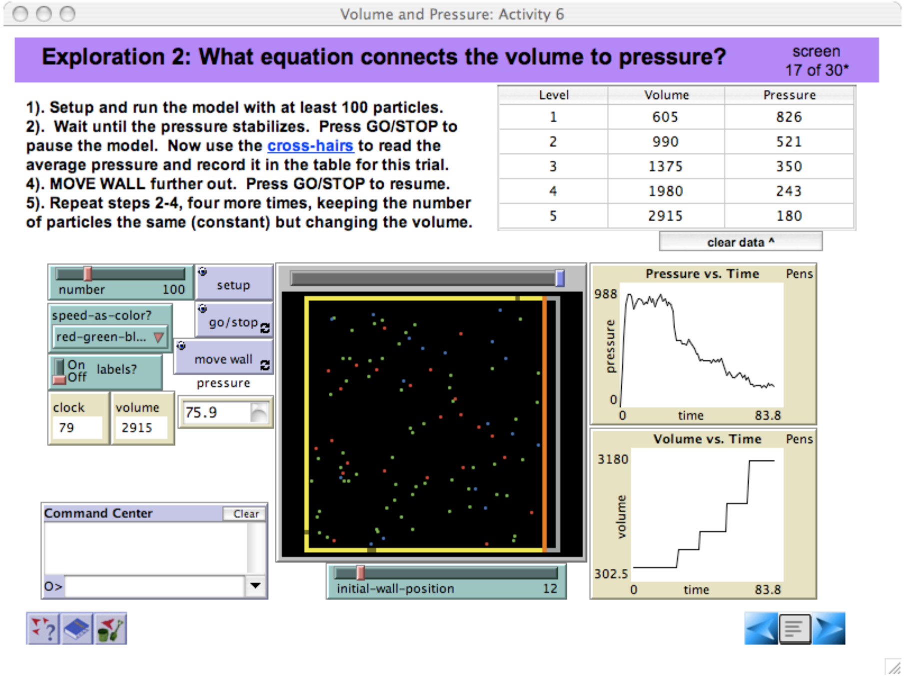
initial-wall-position 12

```
O> ask particle 5 [ pd ]
```

Figure 3: Screenshot from the Volume and Pressure activity in the Connected Chemistry curriculum. Note how the command line is used to instruct a gas particle to leave a path as the volume of the box is increased twice, and the graph, which describes the pressure in the system over time. The question in the script addresses the behavior of a single particle, targeting common misunderstandings of the topic, which can be addressed by observing the microscopic level. Upon submitting an answer, feedback is provided.

Mathematical modeling

Mathematical modeling is prominent in the later activities, where the students are guided in constructing the equations relating the various properties of gases, such as Boyle’s Law (the inverse relationship between volume and pressure, when other variables are held constant). They explore the models, construct scatter plots and compare these to canonical relationships. From these, the symbolic relationships are derived and then used to make and evaluate further predictions. Figure 4 presents a sequence of screens depicting one such activity: the student collects numerical data from the model, relating volume and pressure, observes their scatter plot, compares it with canonical functions and derives the Boyle's Law equation. Further down the line, the students are asked to predict the equations relating the explored relationships, ending up with a construction of the Ideal Gas Law.



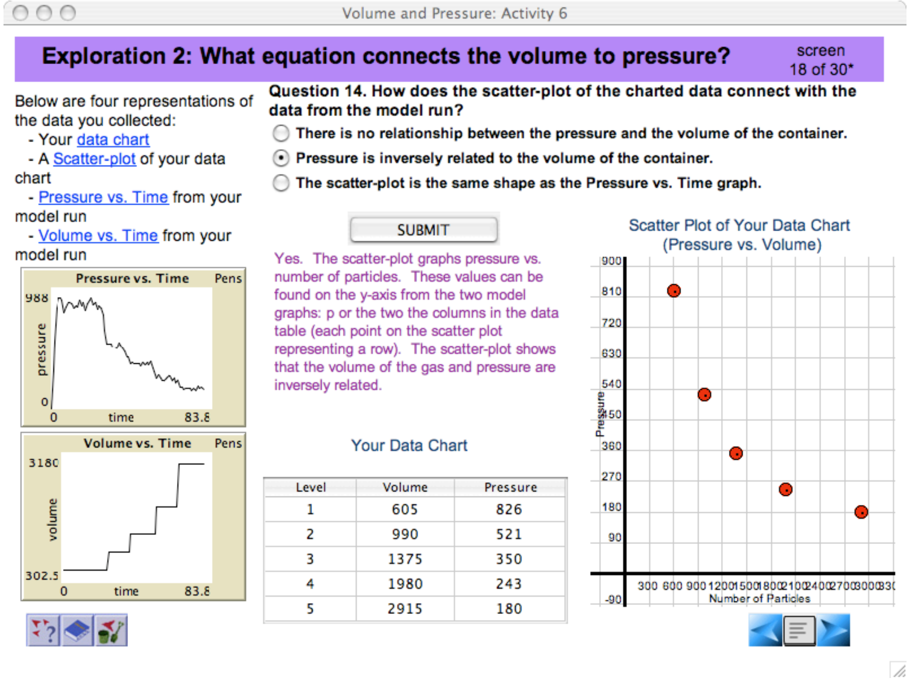
Volume and Pressure: Activity 6

Exploration 2: What equation connects the volume to pressure? screen 17 of 30*

1). Setup and run the model with at least 100 particles.
2). Wait until the pressure stabilizes. Press GO/STOP to pause the model. Now use the cross-hairs to read the average pressure and record it in the table for this trial.
4). MOVE WALL further out. Press GO/STOP to resume.
5). Repeat steps 2-4, four more times, keeping the number of particles the same (constant) but changing the volume.

Level	Volume	Pressure
1	605	826
2	990	521
3	1375	350
4	1980	243
5	2915	180

number 100
speed-as-color? red-green-bl...
labels? Off
clock 79
volume 2915
pressure 75.9
initial-wall-position 12



Volume and Pressure: Activity 6

Exploration 2: What equation connects the volume to pressure? screen 18 of 30*

Below are four representations of the data you collected:
- Your data chart
- A Scatter-plot of your data chart
- Pressure vs. Time from your model run
- Volume vs. Time from your model run

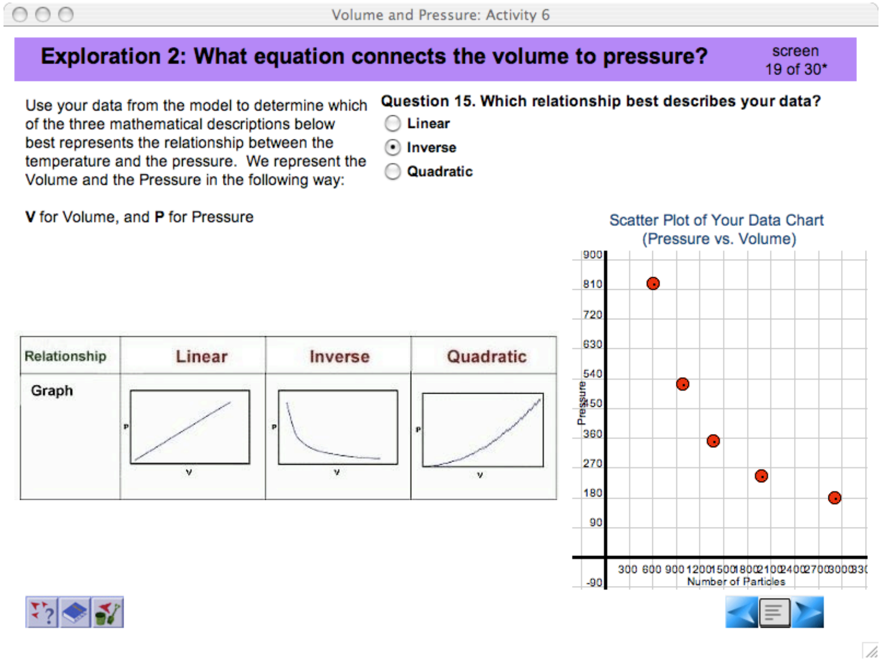
Question 14. How does the scatter-plot of the charted data connect with the data from the model run?

○ **There is no relationship between the pressure and the volume of the container.**
● **Pressure is inversely related to the volume of the container.**
○ **The scatter-plot is the same shape as the Pressure vs. Time graph.**

Yes. The scatter-plot graphs pressure vs. number of particles. These values can be found on the y-axis from the two model graphs: p or the two the columns in the data table (each point on the scatter plot representing a row). The scatter-plot shows that the volume of the gas and pressure are inversely related.

Your Data Chart

Level	Volume	Pressure
1	605	826
2	990	521
3	1375	350
4	1980	243
5	2915	180



Exploration 2: What equation connects the volume to pressure?

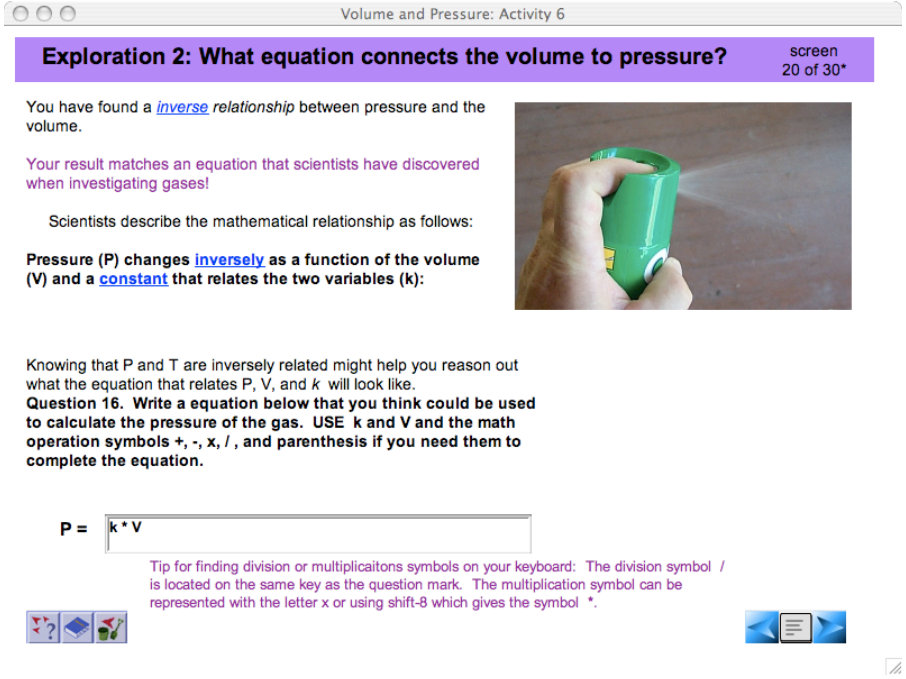
Use your data from the model to determine which of the three mathematical descriptions below best represents the relationship between the temperature and the pressure. We represent the Volume and the Pressure in the following way:

V for Volume, and **P** for Pressure

Question 15. Which relationship best describes your data?

○ **Linear**
● **Inverse**
○ **Quadratic**

Relationship	Linear	Inverse	Quadratic
Graph			



Exploration 2: What equation connects the volume to pressure?

You have found a *inverse relationship* between pressure and the volume.

Your result matches an equation that scientists have discovered when investigating gases!

Scientists describe the mathematical relationship as follows:

Pressure (P) changes inversely as a function of the volume (V) and a constant that relates the two variables (k):

Knowing that P and T are inversely related might help you reason out what the equation that relates P, V, and *k* will look like.
Question 16. Write a equation below that you think could be used to calculate the pressure of the gas. USE k and V and the math operation symbols +, -, x, / , and parenthesis if you need them to complete the equation.

P = k * V

Tip for finding division or multiplicaitons symbols on your keyboard: The division symbol / is located on the same key as the question mark. The multiplication symbol can be represented with the letter x or using shift-8 which gives the symbol *.

Figure 4: Mathematical modeling. Screenshots from the Volume & Pressure activity in the Connected Chemistry curriculum. The student collects data relating volume and pressure, obtains a scatter plot of this data, compares it with canonical functions and derives an equation.

In this paper, we report on the students' learning gains, as they relate to these two goals: micro/macro understanding and mathematical modeling of the system.

Students' exploration of NetLogo models in the Connected Chemistry curriculum

How can we characterize the ways in which students explore computer models to find needed information? Do the types of goals that guide their exploration affect these characteristics? Is the path of exploration affected by the affordances of the model's interface tools? By the kinds of mathematical relationships governing the model's target phenomenon? These questions are explored in a sequence of four studies.

A unique affordance for the research on learning within this environment is the intensive logging of students' actions. One of the exciting opportunities in the Modeling Across the Curriculum project is to virtually "observe" thousands of students as they manipulate models and interact with the embedding scripts. Students' work with the Connected Chemistry models and scripts is collected in logs of their activities, both their text-based activities as well as their actions in manipulating the models and also the model's state and behavior. This intensive logging generates a very large corpus of "click-data", and answers to open-ended and closed questions, for each student. We provide an analysis of the students' model explorations, as derived from computer logs of their actions as they changed parameters and conducted experiments in the Connected Chemistry models. A conjecture of our project is that this data can be mined for features and patterns that reveal important characteristics of the students' exploration and learning. To do so most effectively we are developing automated tools for exploring and extracting patterns from the data. We provide an analysis of students' explorations within computerized models, as derived from computer logs of their actions and the model's changing properties. Through logging the students' actions with the models, we can search for patterns in the students' investigation.

The MAC project consists of several different model-based curricular units. Each of these units was developed independently and can be characterized along a dimension of open-ended-ness. The Connected Chemistry unit is generally the most open-ended of these with many free-form explorations. This presents a particular challenge for analysis of the logs as the students can engage in a wide range of possible actions. In this paper, we focus on the students' exploration of the models themselves.

Framework for studying model exploration patterns and their relation to learning

A framework (see Figure 5) has been constructed to plan the logging and analysis of the data regarding the students' model explorations, as well as their relationship with the students' knowledge and learning. It is based on several key features that make up inquiry in science, as well as a learning and educational perspective:

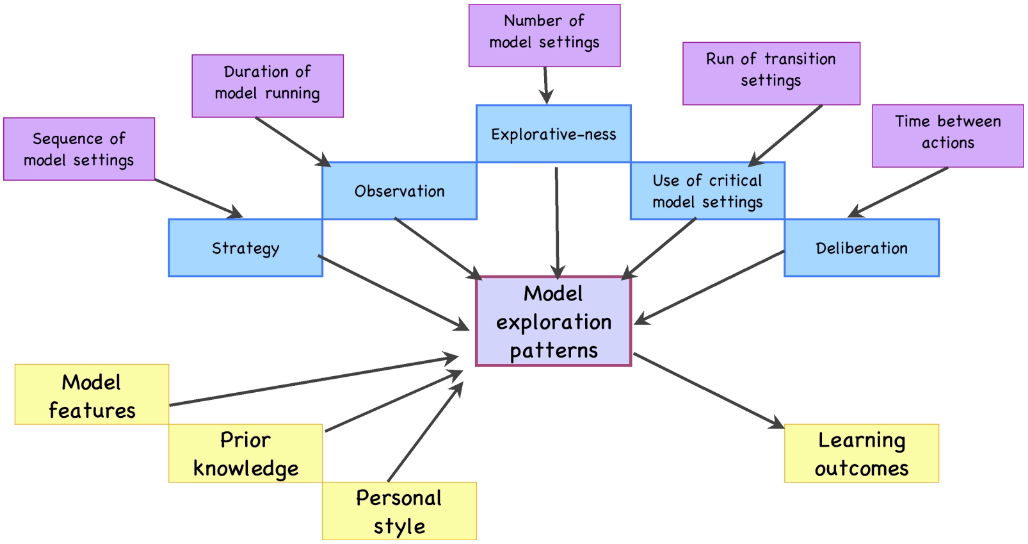


Figure 5: Framework for analyzing students' activities with the Connected Chemistry models, and their relationship to learning outcomes.

Prior knowledge, personal style, and learning outcomes

The anchor and focus of this framework is related to our current investigation: patterns in the students' exploration of computer models. Within the wider agenda of the MAC project, we wish to relate these patterns to the students' prior knowledge, paths of learning and learning outcomes. Prior knowledge may affect the way a model is explored. For example, it is plausible that knowing more about a domain shortens the exploration time, as the student focuses on a few key settings that provide information regarding a specific question. In addition, we assume that personal styles in navigating the model parameter space may impact the way students approach the models in their quest for information. For example, a person who tends to plan ahead and deliberates before taking action will exhibit longer durations between actions. Our framework connects the students' exploration patterns with their resultant learning, or learning outcomes. For example, it is possible that a "click-happy" student, who makes several changes to the model, but spends little time observing its behavior, will not extract enough information from the model exploration to gain a deeper understanding.

What is a student's exploration pattern made up of? In our analysis of these components, we bring in several perspectives: perceptual learning, motor actions, strategies in problem-solving, as well as conceptual issues related to the particular domain and task.

Model running time, observation and perceptual learning

For learning to occur, new information needs to enter the cognitive system and interact with existing knowledge (Samuelson & Smith, 2000). Perception involves the detection and interpretation of sensory stimuli. Perceptual learning is described as a relatively permanent and consistent perceptual change of an array of stimuli, following practice or experience with the array (Gibson, E.J., 1969, 1988, 1991); as relatively long-lasting changes to an organism's perceptual system that improves its ability to respond to its environment (Goldstone, 1998); or – as a discriminating process in which "blurry" impressions are sharpened, differentiated and integrated (Werner, 1957). During learning, perception shifts towards greater correspondence between what is perceived and what is reality.

Observation is a necessary prerequisite for perceptual learning to take place. We cannot assume that if a student spends more time observing a model, these processes of perceptual learning will actually take place; however, it is plausible that when a students spends more time observing the model while it's running, the probability that such learning will occur is increased.

From the logs of the students' model manipulation, we can obtain the duration of a model run. It is probable that the duration of the model runs is related to the time the student actually spent looking at the model. Thus, our gross measure for model observation durations is the time recorded in the logs during which the model is running.

Time between actions and deliberation

Action is defined as motion with intention (Piaget, 1972; Bruner, 1973; Searle, 1981; von Hofsten, 1995), and as such it is distinct from motion alone. Fischer (1980) defines action as the active control of sensorimotor sets, adding the importance of control. The hands serve as channels supporting flow in two directions: enlarging desires into the world (performatory actions) and bringing knowledge from the world (exploratory actions) (Gibson, E.J., 1988; Bruner, 1973; Uzgiris, 1983; McCullough, 1996). While the latter concerns collecting information from the environment, the first is aimed at changing it. In action, we learn the world through feedback from the objects on which we act, so that agreement increases between perception and the world (Frese & Sabini, 1985; Searle, 1981).

The students' actions with the model are exploratory, in that they provide new information about the model's behavior. These actions are organized around distinct functions. For example, exploring the effect of the volume of the box upon the pressure, which the gas particles exert inside the box, involves stopping the model, moving the wall and running the model once more.

In the logs, each action taken in the model is registered. The average time between actions is used as a coarse measure of deliberation. For example, a quick succession of changes to the model could reflect little planning. Longer durations between actions on the model reflect planfulness and deliberation.

Sequence of settings and exploration patterns

A problem can be defined by its conditions: (a) a goal; (b) a barrier that prevents direct access to the goal (Thorndike, 1911, in Rowe, 1987); with Simon (1978, p. 272) adding another condition: (c) attempt or commitment to achieve the goal. Problems can be characterized in different ways: the amount of knowledge needed to solve them (knowledge-poor, knowledge rich, Eysenck & Keane, 1990), the degree to which they are defined (well-defined, ill-defined, wicked, Simon, 1978) and according to the thinking skills that are operated in the process (e.g. Greeno, 1978). Knowledge-rich situations are more difficult to characterize and study because of the amount of knowledge and the variety of ways in which it can be implemented. “Problem-solving strategy” is a term used to describe the way in which an individual chooses a step among all those possible in constructing a solution path towards a target state.

In the Connected Chemistry curriculum, different goals are presented to the students in relatively knowledge-rich problems. One type of goal is discovering *qualitative relationships* between variables, such as volume and pressure, in terms of “more”, “less”, “increase” and “decrease”. Another type of goal involves noticing *distinct model behaviors in a particular regime* of the parameter space. For example, in exploring “how is pressure determined in the model?”, the student would benefit by setting a small number of particles in the box and connecting between their hitting the wall with zero and non-zero values in the pressure monitor. Another type of activity involves collecting data in order to derive a *quantitative relationship* (e.g., Boyle’s law). Each of these activities is framed by a different goal. Different goals may encourage different strategies in exploring the model. The open-ended form of some of these problems makes them ill defined; other problems are more highly structured, and would be termed “well-defined” problems.

We examine the impact of the different goals on the students’ exploration patterns. In the logs, each new state of the model following a change in the model is recorded, e.g. the box volume set by the student as she investigates the relationship between volume and pressure. We examine these states as a sequence of settings the students employ in their exploration. We notice their order and their relative magnitude along a time-line of the students’ investigation.

In this paper, we report on the how we typify such sequences and demonstrate students’ explorations into a particular topic.

Domain specific features of model exploration

In the more general domain of complex systems, as well as the specific topic of the complex gas particles system explored in the Connected Chemistry curriculum, additional aspects of the exploration become important.

We incorporate two features in the framework, which are related to complex systems. One is the richness of the exploration, as reflected in the number of different settings, which a student employs. The other is the use of critical settings.

The behavior of a complex system is not linear. For example, in a rigid box when more and more particles are added (or pumped in), the system does not respond in similar ways to different additions. When there are few particles in the box, they are virtually independent of each other, each colliding with the wall, barely colliding with each other. In this regime of the parameter space, one can say that the “whole is the sum of its parts”. However, beyond a certain density, or critical value, the collisions or interactions among the particles become more dominant. At this point, the speeds and paths of the particles are not determined solely by the box and their own properties; but also by their energy-and-momentum-conserving interactions with other particles in the box. We can see the distribution of particles’ speeds in the box as reflecting such non-linear behavior. When many more particles are added in the box, we can see additional departures from previous model behaviors. At some point, the collisions become so dominant, that a single particle may be “trapped” in a smaller section in space. At this point, a "division of labor" among the particles emerges. Some particles are close to the wall and keep hitting it repeatedly, raising the pressure. Other particles seldom reach the wall and do not contribute to the overall group pressure.

While we explicitly incorporate only some of these principles in the curriculum, they are all "out there" in the models and can be explored by the students. Several strategies can benefit by noticing these features of the model. For example, moving in small intervals through the parameter space can expose the points at which the model departs from one behavior to another.

In capturing the students' model settings, we can see how many runs were made. By looking into their specific values, we can discover whether different regimes were accessed. In the current paper, we analyze only the number of different runs. We intend to investigate the behavioral regimes and critical settings in a future analysis.

Previous work

In previous work, we have explored students' exploration patterns within different types of activities, using different tools and their relation to prior knowledge (Levy & Wilensky, 2005, 2006a, 2006b). This work resulted several conclusions:

Patterns in exploring models

In a relatively open task where the goal state is a narrow range of values, students displayed three distinct patterns regarding the succession of settings they employ: "straight to the point", "inching through space" and "oscillating". The "straight to the point" pattern is one in which the actions are spaced out in time, and each run is observed for longer times, even though the number of runs is very small. The "inching through space" pattern approaches the goal state in steps, each succeeding step smaller and closer to the target, displaying convergence. The actions are very close together in time, the student spends little time observing the results of each run, even though the number of runs is greater. The "oscillating" pattern describes moving back-and-forth between the target state and its opposing pole in the parameter space. It is associated with closely spaced actions, short observation time per run, which is complemented by longer overall observation time. This study was conducted with a small data-set, and is continued in this study with a larger sample.

Goals and exploration patterns

Students explore the models in characteristic ways across tasks with different goals. Model exploration when there are different goals is distinct in one aspect: when there is a particular goal state the students gravitate towards this state; when a range of values is informative, the students span a wider range of values. However, across tasks, most of the students were consistent in which pattern they used to explore the models.

In activities aimed at deriving a quantitative relationship, some of the students employed a new strategy: "constant intervals", where the independent variable in the experiment is increased at constant intervals. This pattern was not seen in open activities, which were aimed at qualitative relationships.

Tool affordances and exploration patterns

When varying settings via textual NetLogo commands, rather than through a linear slider, a large proportion of the students used an oscillating pattern, shifting back and forth across orders of magnitude, rather than the more limited space offered via a slider. The greater freedom this tool offers afforded their exploration along a greater range in the model's behavior space. Not only did they explore a larger range, the sequence of values was not linear.

Mathematical nature of the model and its exploration

In a highly scaffolded model with a linear function underlying the model's behavior, almost all the students used the "constant intervals" strategy. However, even then, a large portion of the students used alternative strategies. In a less scaffolded model with an inverse function underlying the model's behavior, almost half of the students adapted their exploration to the inverse function and a wider distribution of strategies was observed. Prior knowledge is associated with the strategy the students use in exploring the three models.

Focus

In this study we continue this line research, both with larger data-sets and with an improved analysis scheme. We delve deeply into one student's progression. We compare the main goals of the Connected Chemistry curriculum with students' learning gains.

Based on our theoretical framework and previous work, three goals guide this investigation:

(a) Expanding the coding scheme used to characterize students' model exploration strategies in open-ended explorations.
(b) Examining one student's progression from beginning to end, highlighting the affordances of logging to mine valuable information.
(c) Describing students' learning gains resulting from activity with the curriculum in terms of gaining a micro/macro perspective and mathematical modeling.

METHOD

Sample

As part of the Modeling Across the Curriculum project, during the 2005-6 school year, the Connected Chemistry curriculum has been implemented in nine schools, 27 classes.

For the study into students' exploration patterns, 123 students were randomly selected from among three schools that had engaged in the Connected Chemistry activities earlier in the year. For the study into students' learning gains, 605 students completed both pre-test and post-test and are included in the sample.

Data collection and analysis

Students' exploration patterns

We have gathered a large corpus of data, recording students' responses to both multiple-choice and open-ended questions, as well as student "gestures" as they interact with the computer models. In the first part of the paper, we focus on the latter – what characterizes the students' exploration of the Connected Chemistry models, themselves.

In logging the students' activities in the Connected Chemistry environment, we collect each action the student takes: multiple-choice and free-text answers, a well the NetLogo models manipulation. Thus, we have information on the following: (a) initial settings; (b) pressing and un-pressing a button; (c) change in a slider or a switch; (d) entering a NetLogo command in the Command Center; (e) Slowing down the model; (f) states of the model when any action is taken.

We investigate the students' exploration patterns in two focal screens (see Figure 6) in one of the early activities: "Changing pressure". The activity engages with the following idea: the pressure of a gas in a container (macro property) is related to the gas particles hitting a surface, the walls in the container (micro behavior). One screen that introduces the model and some of its new features precedes the focal screens. In the focal screens, the students are asked to make the pressure monitor read zero. Possible solutions are either having no particles in the container, so that the pressure is always zero; or having very few particles in the box, so that the pressure sometimes reads zero. Thus the target in the parameter space is a small number of particles.

In the model, the students can change only one setting: the number of gas particles in the box. We classify this activity as relatively open, as the students can select any value for this setting. Thus, the feature we focus on here is the sequence of values they set in successive runs of the model for the number of particles.

In this analysis, we extracted the students' successive settings in running the model. These are portrayed as temporal graphs of the settings (see Figure 7). The process of obtaining the information in these studies was performed in a combination of automatic and manual modes. The sequence of settings the students used in the focal screens was mined automatically as a list of settings for each student. The parser detects the student's entry into the focus screens, and extracts the distinct settings that have been selected and run. Analysis of the list in terms of exploration patterns was done manually. We are currently in the final stages of automating the process. Due to the manual mode of analysis, our sample consists of 123 students. Once full automation is achieved, we will be able to rapidly analyze the large data samples we have collected.

In preparing for an automated calculation of model exploration patterns with respect to the sequence of settings, we decided to shift away from the clustered and complete patterns we had previously detected (Levy & Wilensky, 2005, 2006a). Our intention was to step away from possibly misconceived patterns, and examine their features with a larger sample before attempting to re-cluster them. We converted these patterns into dimensions:

1. Oscillation – number of direction changes (detects "oscillating" pattern)

2. Step size – with cutoff value, above which these are “big” steps, and below it these are “small steps” (detects “inching through space” pattern). The cutoff value was determined empirically by creating a histogram of the students’ step sizes.
3. Path length - how many settings until target zone is reached (detects “direct to the point” pattern)

The students’ sequences of settings were coded for the three dimensions.

In order to fully understand students' interaction patterns, we have been analyzing logs of students' navigation through the activities, interactions with the graphical interface, and NetLogo commands that they typed. The logs, formatted using XML, were written to disk as the student finished each activity, and were transmitted to a central server soon afterwards. Our data thus consists of more than 12,000 logfiles, each representing one student's exploration of a particular activity. Because students may have explored an activity during more than one class session, it is possible for a particular student-activity combination to exist in more than one logfile.

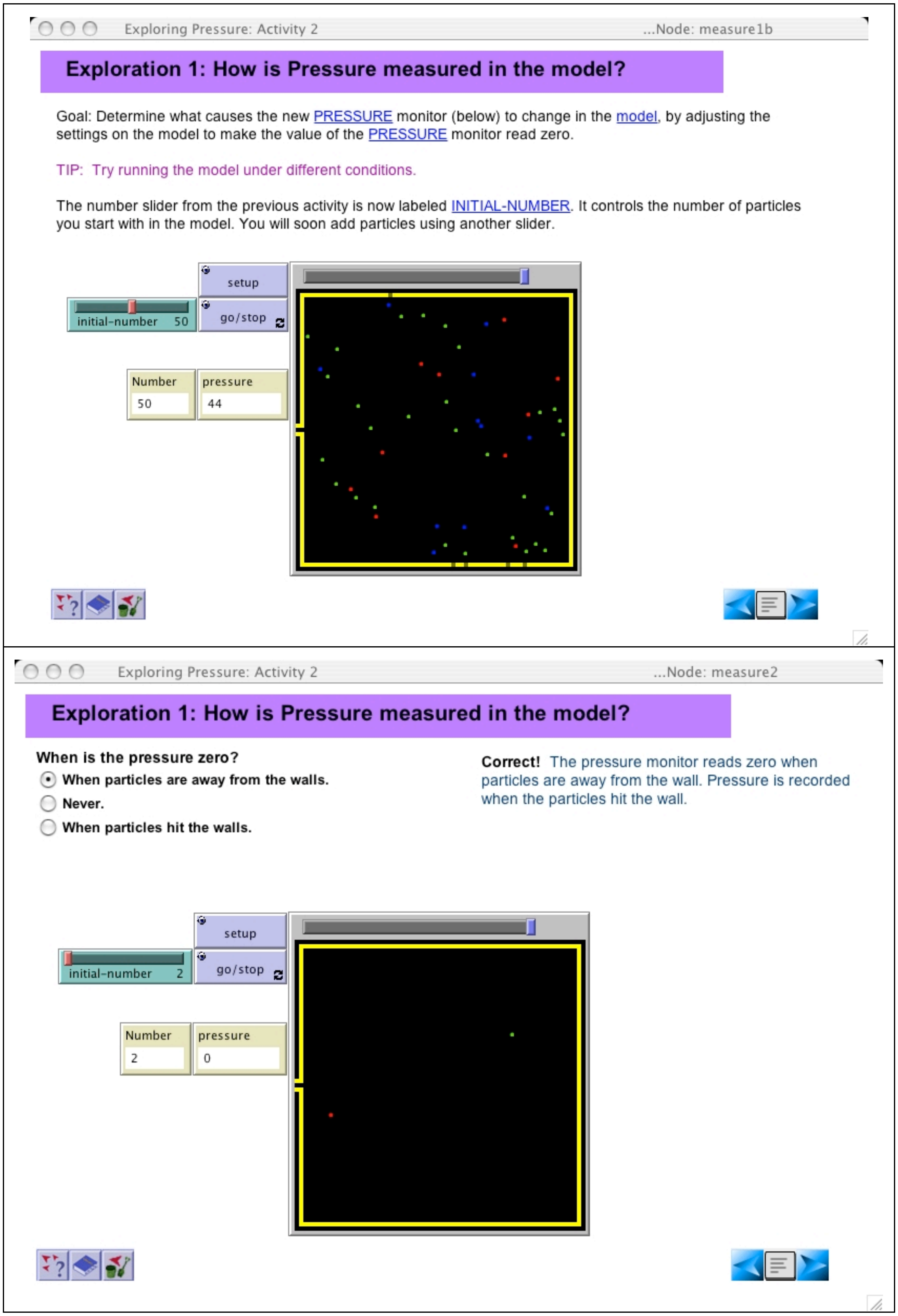
Initial analysis of the original XML files proved slow and difficult. The data was thus imported into a series of normalized tables in a relational database (RDBMS). For certain types of analysis, such as determining the amount of time students spent in particular activities and models, the database was an excellent choice, providing our team with an easy and efficient access to the data.

By contrast, finding our desired patterns has been quite challenging. Our initial queries, against fully normalized tables (Celko, 1999) resulted in extremely long-running queries, some of which would have taken weeks to complete. We then created a centralized "fact table" (Kimball and Ross, 2002), adding a number of columns for precomputed values, in order to speed up later retrieval queries. However, this still did not speed up the queries sufficiently, and left open the question of how to describe, identify, and retrieve queries matching the students' interaction patterns.

We turned to the discipline of data mining (Han and Kamber, 2006) for assistance. While data mining often has the goal of finding unexpected patterns, we used a number of data-mining techniques -- specifically, data cleaning, transformation, and reduction -- to improve the performance of our database and the accuracy of our pattern-matching programs. Segmenting the fact table into separate databases, each of which contained logfile data for a single activity, also helped to improve performance, and also reduced the complexity of many queries.

Despite all of these steps, querying the database for specific patterns has remained difficult and slow going. Upon further examination of the data-mining literature, it would seem that our data is significantly more complex than the typical data fed to a data-mining system, and that the patterns for which we are looking are similarly problematic. Specifically, the fact that we are looking for interactions among multiple variables, that the data cannot be classified as even approximately periodic, and that many values are recorded several times per second have made the queries difficult to construct, as well as slow to execute. Added to this is the fact that some of the most-important patterns are contextual, meaning that they require a comparison between multiple lines in the database. Such temporal, non-periodic patterns are difficult to describe and slow to execute (Han and Kamber, 2006; Snodgrass, 2000). The fact that variable values are repeatedly logged, even when they are unchanging from the previously logged value. Finally, retrieving every row of a database in order to compare it with other rows might be the only way to produce results -- but it is a significantly different use case than the database authors intended, and one for which the database is not optimized.

We are working toward a multi-pronged solution for future pattern-matching work: We are trying to obtain faster hardware, and are also looking into algorithms that scale up more efficiently. We are also looking into alternative ways to represent the data, which might open the door to faster and more established techniques. For example, if we can find a way to represent the data warehouse as a text string, we would undoubtedly be able to take advantage of existing regular-expression engines (Thompson, K.). Work in bio-informatics, in which researchers compare gene samples with known standards (or averages) might also provide useful solutions or insights.



Exploration 1: How is Pressure measured in the model?

Goal: Determine what causes the new PRESSURE monitor (below) to change in the model, by adjusting the settings on the model to make the value of the PRESSURE monitor read zero.

TIP: Try running the model under different conditions.

The number slider from the previous activity is now labeled INITIAL-NUMBER. It controls the number of particles you start with in the model. You will soon add particles using another slider.

initial-number 50

Number 50

pressure 44

Exploration 1: How is Pressure measured in the model?

When is the pressure zero?

● **When particles are away from the walls.**
○ **Never.**
○ **When particles hit the walls.**

Correct! The pressure monitor reads zero when particles are away from the wall. Pressure is recorded when the particles hit the wall.

initial-number 2

Number 2

pressure 0

Figure 6: Connected Chemistry “Changing pressure” screen-shots in the analyzed activity

Learning gains

We created a pre-test and post-test questionnaire with 18 multiple-choice items, which targets the main concepts and some of the skills, which we aim to develop through the Connected Chemistry curriculum. Since many of the skills, such as model exploration patterns, can be assessed directly via logging of the actions in the activities, we did not include these in the pre-test and post-test.

The questionnaire was planned by using a two-dimensional analysis of the entire curriculum[1]. This analysis is based on the main content topics, or the learning objectives, as one dimension, and upon Shavelson's knowledge dimensions (Ayala, Shavelson, & Yin, 2002; Shavelson, Li, Ruiz-Primo, & Ayala, 2002). The questions that were asked of the students during the activities were coded in the following way:

(1) Learning objectives: Kinetic Molecular Theory and gas particle behaviors; Content specific to the complex systems perspective – micro-macro relationships and processes of perturbation and equilibration; Gas Laws.
(2) Forms of knowledge: declarative, procedural, schematic and strategic.

After calculating the proportion of questions in each dimension, the same proportions were used to plan the questionnaire. This is true for all but one category: the procedural form of knowledge. We have included a smaller proportion of this category, as direct analysis of this knowledge via the activities is possible via logging, and in the interest of brevity.

Some of the items were selected from research studies that targeted specific misconceptions in the field. Items regarding Boyle's Law were previously developed and used by Bowen & Bunce (1997). Most of the items were invented "in-house" to reflect both the topics we relate to in the activities, and specific confusions we have detected among students in our design-based research of the activities.

In the process of writing the questionnaire, the draft underwent several reviews by the team at the Center for Connected Learning and Computer-based Modeling, as well as researchers from the Concord Consortium. It was revised twice in accord with the development of new activities and lessons learned from analysis of the previous versions. For example, compound questions were broken down to detect the source of error in understanding.

The questionnaires were then scripted and made available to the students.

We measured the 2005-6 students' learning gains using the pre- and post-tests. We analyzed the results from the students' answers in terms of the following dimensions:

1. Overall
2. Chemistry/Complex systems content related:
    2.1. Items that focus on micro-level understanding: Understanding of the system's behavior at the particle level, e.g. understanding that when particles hit the wall of a container, they change their direction but not their speed.
    2.2. Items that focus on macro-level understanding: Understanding qualitative

---

[1] A similar scheme was previously used for the Biologica™ and Dynamica™ curricula in the MAC project.

relationships regarding the system's population level variables, e.g., Boyle's law, the inverse relationship between volume and pressure.

  2.3. Items that focus on connecting micro and macro levels: Connecting the two description levels, e.g., understanding that a rise in pressure in a fixed container is related to the particles colliding more often with the walls of the container.

3. Mathematical modeling skills related:
   3.1. Solving numerical problems: Solving quantitative problems, e.g. calculating the pressure in a piston, when its volume is changed.
   3.2. Graph interpretation: Understanding how a graph describing an aspect of the system's behavior is related to phenomena in the real world, e.g. A graph of pressure in a basketball rising suddenly at some point in time can be associated with someone sitting on the ball, but not with putting it in the freezer.
   3.3. Deriving equation from graph: Translating a graph into an equation form; e.g. selecting an equation that describes the relationship seen in a graph of Volume vs. Temperature.

RESULTS

We present the results in three sections: (1) Students' exploration patterns; (2) A case study with one student; (3) Learning gains.

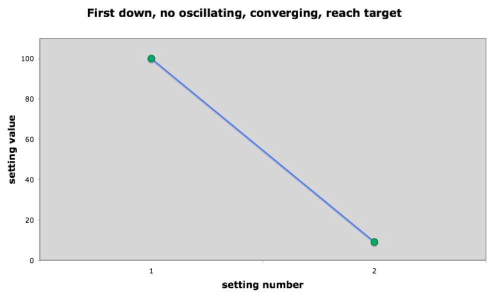
Students' exploration patterns

The students' sequences of settings were coded for the three dimensions. The results are displayed in the following Table 1 and Figure 7:

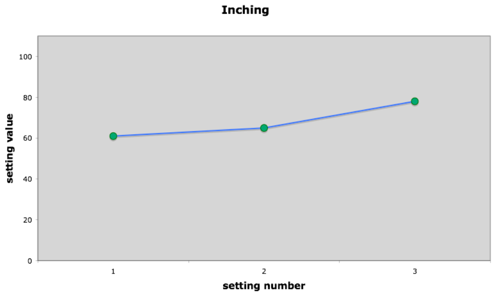
Dimensions	Proportion of sample
Straight	12%
Inching	11%
Oscillating	28%
Straight & Inching	3%
Straight & Oscillating	10%
Inching & Oscillating	10%
All three	11%
None	15%

Table 1: Dimensions of exploration among 123 students manipulating a model in the Changing Pressure activity in the Connected Chemistry curriculum.

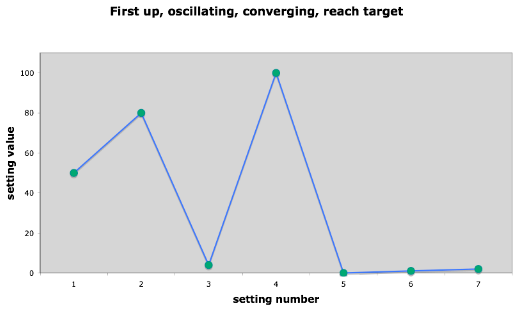
Straight to the point (12%)



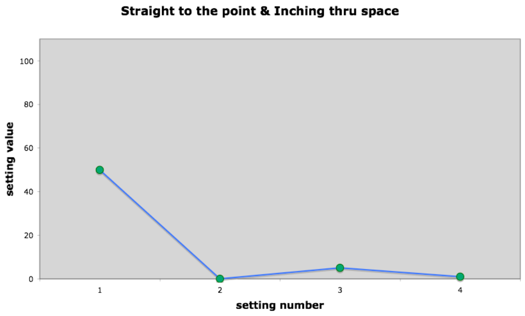
Inching through space (11%)



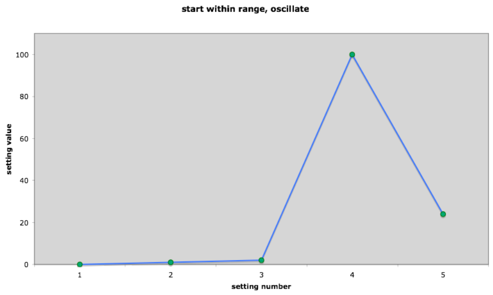
Oscillating (28%)



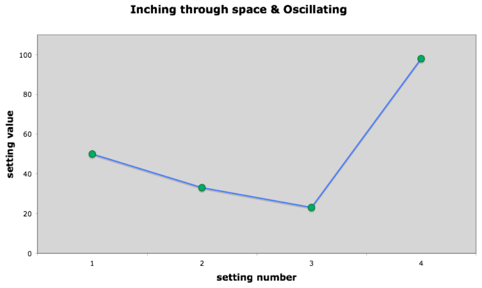
Straight to the point + Inching (3%)



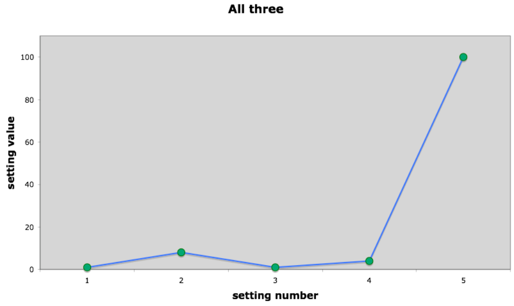
Straight to the point + Oscillating (10%)



Inching + Oscillating (10%)



All three (11%)



None (15%)

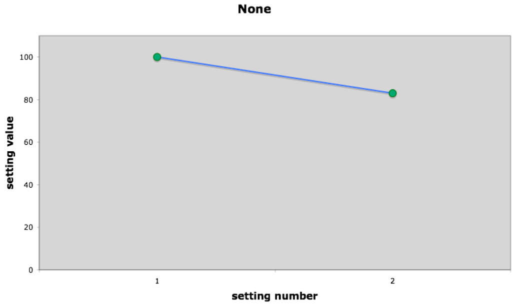


Figure 7: Canonical examples of students' exploration patterns while manipulating a model in the Changing Pressure activity in the Connected Chemistry curriculum. Y-axis describes the settings they use, the number of particles. X-axis is the setting order.

We conclude:

1. 51% of the patterns are "pure" in the sense that they reflect only one of three original patterns (Levy & Wilensky, 2005). 15% do not fall into any of the patterns: these are students that never reached the target zone. The rest, 34% show a combination of these patterns. The most dominant pattern is "oscillating" between extremes. The rest are evenly distributed, apart from the "straight & inching" that is relatively infrequent.
2. As seen in the canonical examples, the students that exhibit combined patterns usually (24%) display two exploration phases: in the first phase, the target zone is reached (small number of particles); after figuring out the answer to the guiding question, they do not leave the screen, but continue exploring the parameter space – either moving between extreme values (oscillating dimension) or moving gradually through space (inching dimension). Thus, the students flexibly shift among these exploration patterns as their goals for the activity move from externally to internally directed explorations.

A case study - Student number 4981

In this section, we illustrate the path of one student through the curriculum, highlighting the student's typical and atypical manipulation of the models. From describing the pre-test and post-test results for this student, we delve into the activities and attempt to understand some of the observed shifts. The case study demonstrates the affordances of intensive logging for gaining insight into students' learning.

Student number 4981 was selected randomly out of a large pool of data. We do know that she is a 9th grade female from a medium-sized high-school, with a highly diverse population, speaking 16 languages apart from English.

When observing her pre-test and post-test answers we note the following. Walking in, this student is quite knowledgeable regarding several aspects of gas behavior: She has a correct mental image of the gas particles' spatial distribution and she knows the canonical definition of pressure. Regarding the various relationships among the aggregate gas variables, we observe that qualitatively – she had a good grasp of these relationships; however, quantitatively, she can solve correctly only for Boyle's Law. Her ability to reason with graphs, which depict a physical situation, is limited as well. She is not acquainted with the gas particles' behaviors, or with the KMT (Kinetic Molecular Theory) assumptions regarding changes in particles' direction and speed upon collision. She assigns intentionality to the gas particles as they move into a vacuum.

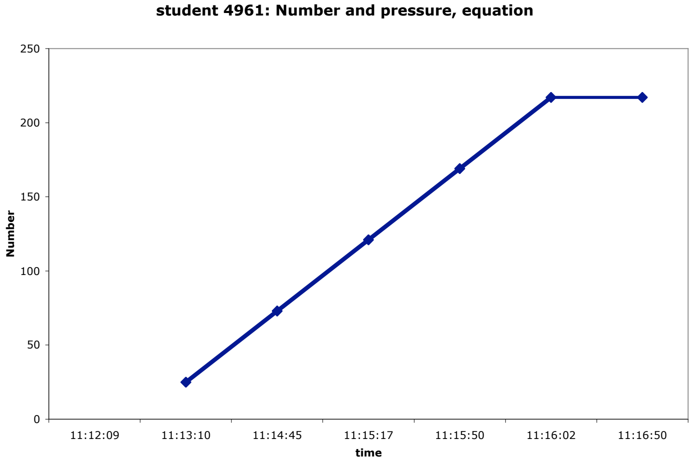
In the post-test, we can see some domains of improvement. Student 4981 improved regarding her understanding of KMT and the individual particle behaviors and in the quantitative aspects of problem-solving in the domain: understanding and implementing the relationship between aggregate system properties, interpreting the graph representation of their temporal changes and deriving an equation from a graph. However, she still assumes that particles have intentions, and has difficulties relating micro-particle-behaviors to aggregate properties of the system.

What happened between the pre-test and the post-test? We base the following on the logs of her answers to the questions in the script and from the way she manipulates the models.

This student quickly grasps the tools and representations, and has only few problems in manipulating the models. We have not observed any typical errors students had made in manipulating the models, which resulted from not following through the accompanying instructions[2].

In each of the seven activities' post-assessment, she is completely successful. Thus, within the activities, we can see greater learning gains that are not necessarily evident in the post-test results. In the activities, she can tie the micro-to-macro relations, in a way that was not seen a few days later in the post-test.

We note that overall, Student 4981 does not explore the models beyond the tasks' minimum requirements. She uses the models in a perfunctory fashion, according to the scripts' suggestions and no more. For example, when the script suggests that she change the number of particles in the container and observe the resultant pressure, she does this – but only once. We have seen other students employ several settings in the very same activity. In the activities geared at deriving an equation (see Figure 4), we have seen her use mainly a "constant intervals" pattern of exploration (See Figure 8). This, together with the pre-test results and her carefully following the scripts' instructions demonstrate that this student is well adapted to the classroom environment.



---

[2] An example of a typical error is changing initial variables after the model has started running. We have addressed such problems by adding several short intervening activities to scaffold the students' manipulation of the models.

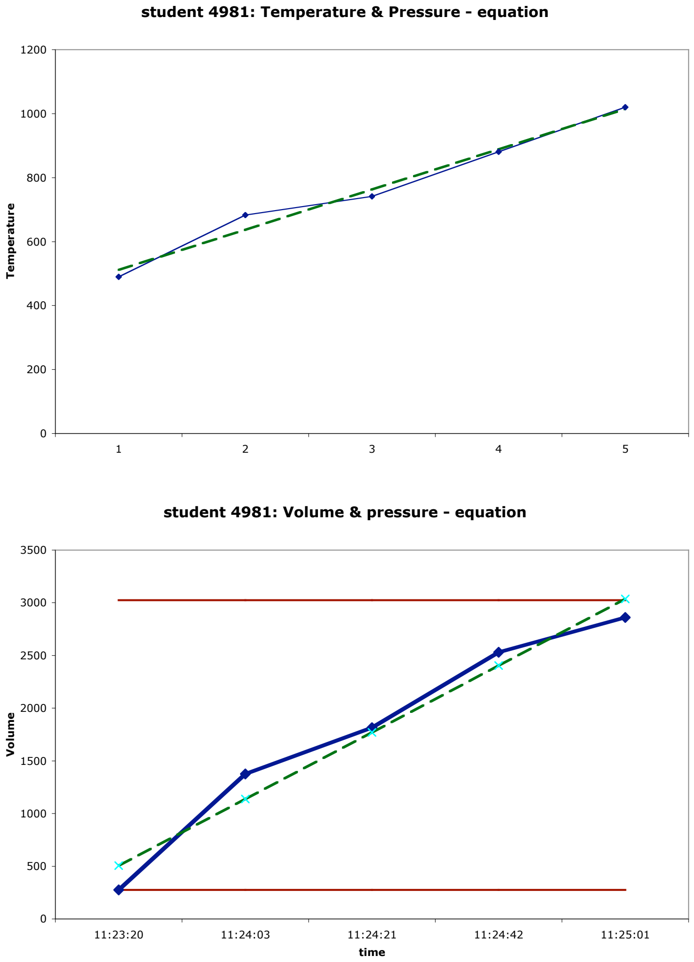


Figure 8: Student 4981's explorations of the models, geared at deriving an equation relating aggregate variables for the behavior of a gas. The green dotted line represents the closest fit of a "constant intervals" pattern. For the "Volume & Pressure" exploration, a 5% deviation from the fitted function is still considered "constant intervals" since the change in volume is made via clicking on the model, rather than changing a variable slider.

In most cases, we do not see any playful investigations or inventive patterns of exploration. Thus, it is interesting to note when she departs from this pattern. This happened in two instances. Both instances relate to shifts among micro- and macro-levels, which she found difficult in both pre-test and post-test.

In the second activity "Changing pressure", she investigates "When is pressure zero?" (for details, see previous section on students' exploration patterns) during 2:13 minutes. This activity targets the more minute interpretation of pressure in a micro-to-macro approach, highlighting the distinction among collisions taking place among particles in two dimensions, and the one-dimensional wall-hits. We can see the sequence of settings she employs for the number of particles in the box in Figure 9: 50, 53, 69, 36, 4, 1. She starts out with the model's default value of 50 particles, gradually raising the value to 69 particles. After this, she begins to "slide" down to a low value for this number, in a pattern we have called "inching through space". The intervals among settings become smaller and smaller, as the target region is approached, ending with a single particle in the box. We can interpret this sequence in the following way. Initially, she increases the number of particles twice, possibly reflecting an alternate understanding of the system, or having no clear idea of which way to go with the exploration. However, observing the resultant pressure clues her into a more fruitful path of investigation: go down. She reduces the number of particles, but not enough – the pressure is lower, but never zero. Two jumps down, and she reaches the regime where the pressure is *sometimes* zero – four particles and then one. With such a small number of particles in the box, the pressure is mainly zero – fluctuating up when the particles hit the container's walls. We note two important points. This activity targets a connection between individual particle behaviors and the aggregate system properties, an aspect she has not demonstrated an understanding for. This invites her intense exploration, examining the system along an intermediate range, until this connection is clarified. Once this is settled, the exploration ends. We also have data on other features of the way she manipulates the model: The model was actually running during 1:13 minutes, or an average of 0:12 per run, a short duration with respect to other students (Levy & Wilensky, 2005). She has performed 22 actions upon the model, an average of 6 seconds between actions. This is relatively fast, quite distinct from her other explorations. To summarize this activity, the student has shifted from a minimal yet planful style of exploration to another strategy "inching through space". This happened following a good question and subsequent confusion.

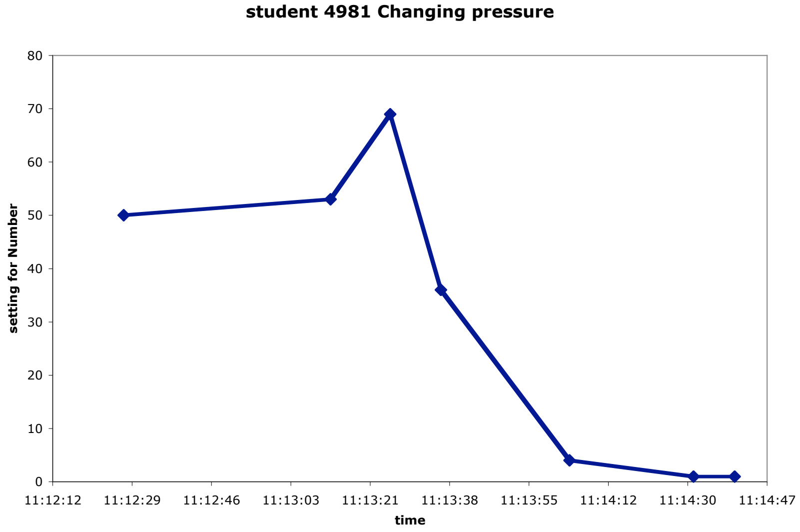
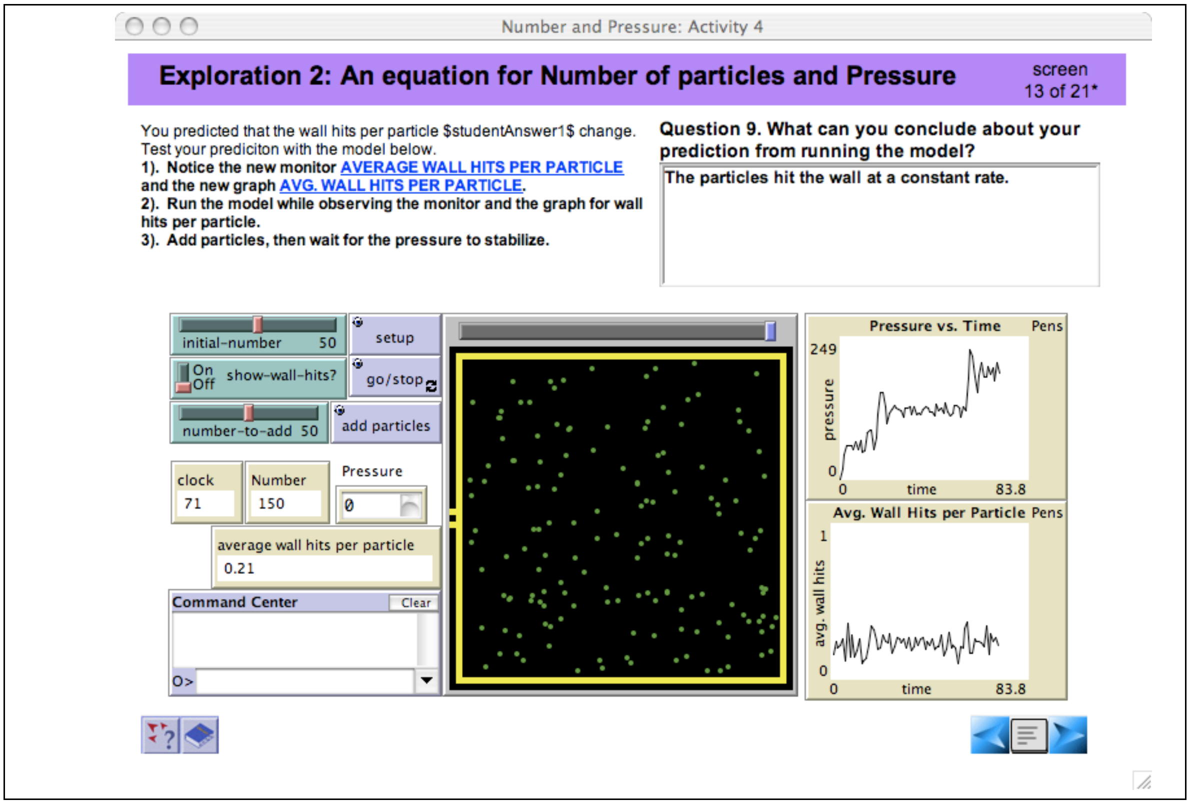


Figure 9: Student 4981's exploration of "When is the pressure zero?" in the "Changing pressure" activity.

Another puzzle is posed in the fourth activity "Number and Pressure". While the pressure increases linearly with the number of particles in a container, the average rate at which a *single* particle hits the wall remains *constant*. This is another case in which we highlight the distinction among the micro and macro levels of description in the system (see Figure 10). In predicting what will happen before testing this out in the model, she has predicted: *"By rate as they hit the wall mean the force they exert on it? Or does this mean the rate (how often) each individual particle hits the wall? I'm not sure about the question. But I think that it will increase, although I do not know the* *rate."* We can see her confusion regarding this question, is it forces they exert? Or rate of hitting the wall?: *"I'm not sure about the question".* As most students, she finally predicts that the rate will increase. Her forthcoming exploration is highly intensive (see Figure 11). Twenty settings are employed during almost two minutes of exploration over an extremely large range in the parameter space. Once again she has turned to an "inching through space" strategy of exploration. The puzzle elicits a very different style of manipulating the model – closely spaced multiple settings, as a wide range of the parameter space is employed. We note the similarity in strategies with the previous exploration, and their distinction from the perfunctory investigations in the rest of the curriculum. In answering the question immediately following the exploration, she types (the scripts' question are bolded): [**What can you conclude about your prediction from running the model?**] *"From this model, it doesn[']t seem to matter how many particles there are. The average number of wall hits per particle is still around the same."* [**The average number of wall hits per particle does not change significantly with more particles. Why?**] *"I think that no matter how many particles there are, each particle still has the opportunity to collide with the wall. I do not think density with the wall matters."* She has ended at a very different place from which she had started. She had started with a merged view of description levels – "if the aggregate rate at which particles hit the wall increases, so does the individual rate for each particle". She ended with a distinction among the micro- and macro-levels, describing the system.



Number and Pressure: Activity 4

Exploration 2: An equation for Number of particles and Pressure screen 13 of 21*

You predicted that the wall hits per particle $studentAnswer1$ change.
Test your prediciton with the model below.
1). Notice the new monitor AVERAGE WALL HITS PER PARTICLE and the new graph AVG. WALL HITS PER PARTICLE.
2). Run the model while observing the monitor and the graph for wall hits per particle.
3). Add particles, then wait for the pressure to stabilize.

Question 9. What can you conclude about your prediction from running the model?

The particles hit the wall at a constant rate.

initial-number 50
show-wall-hits? Off
number-to-add 50
clock 71
Number 150
Pressure 0
average wall hits per particle 0.21

Figure 10: Connected Chemistry "number & pressure" activity – a puzzle regarding the overall increase in rate at which the gas particles hit the wall upon adding particles into the container coupled with a constant rate at which each particle hits the wall.

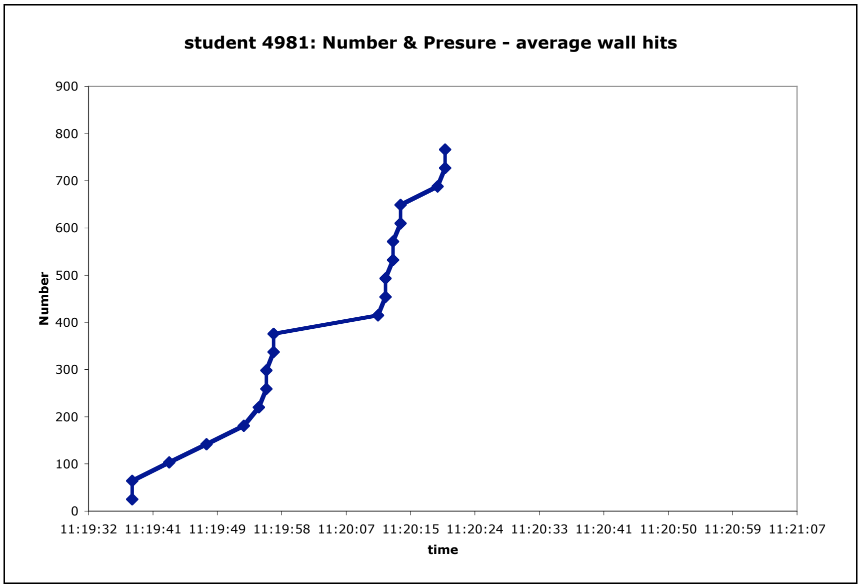


Figure 11: Student 4981's exploration of the activity highlighting the distinction among the increase in pressure and the constant rate at which each particle hits the wall in the "Number and pressure" activity.

Learning gains

The following results were obtained for the students' pre-test and post-test scores:

	Measure (max score)	Pre-test score	Post-test score	t	Sig (2-tailed)
	Total score (18)	8.89	10.34	-11.48	< 0.001
Chemistry/Complex systems content	Micro (3)	1.2	1.8	-13.11	< 0.001
	Macro (6)	3.5	3.8	-4.58	< 0.001
	Micro-macro (3)	1.2	1.6	-7.85	< 0.001
Mathematical modeling skills	Solving numerical problems (2)	1.2	1.3	-1.89	0.059
	Graph interpretation (3)	1.2	1.5	-7.48	< 0.001
	Deriving equation from graph (1)	0.38	0.54	-6.67	< 0.001

Table 2: 2005-6 pre-test/post-test scores, broken down by dimensions of learning (N = 605)

For five out of six dimensions as well as in the total score, students made significant gains. The largest gains were in understanding chemical phenomena at the micro- level and their connection to the macro-level phenomena as well as graph interpretation and constructing a relational equation from scatter-plot data. Regarding numerical problem-solving we have seen non-significant improvement.

DISCUSSION

Our investigation was guided by three goals:

(a) Expanding the coding scheme for characterizing students' model exploration strategies in open explorations targeting the understanding of a qualitative relationship.

(b) Examining one student's progression from beginning to end, highlighting the affordances of logging to access valuable information.

(c) Describing students' learning gains resulting from activity with the curriculum in terms of gaining a micro/macro perspective and mathematical modeling.

We turn to discuss the results with respect to these goals.

Students' exploration patterns: refining the coding scheme

How do students search for information within computer models? Previously, we have found that activity goals, available tools, underlying model behaviors and personal styles interact in shaping the particular form by which information is searched for (Levy & Wilensky, 2005, 2006a, 2006b). These forms were described via a multi-faceted framework (Figure 5), which incorporates perceptual learning, motor actions, problem solving and domain-specific features. This framework is situated within a wider structure that seeks the relationships between prior knowledge, learning paths through interaction with computer models and learning outcomes. In this paper, we focus upon the students' activity with the models, their selected sequence of settings.

In our previous work, using a smaller data-set, we had described students' exploration patterns in terms of sequence settings, action rate, observation time and the number of model runs. We identified and described three prominent patterns: "straight to the point", "oscillation between extremes" and "inching through space". The "straight to the point" pattern is one in which the actions are spaced out in time, and each run is observed for longer times, even though the number of runs is very small. The "inching through space" pattern approaches the goal state in steps, each succeeding step smaller and closer to the target. The actions are very close together in time, the student spends little time observing the results of each run, even though the number of runs is greater. The "oscillating" pattern describes moving back-and-forth between the target state and its opposing pole in the parameter space. It is associated with closely spaced actions, short observation time, which is complemented by longer overall observation time.

In the current study we have re-conceptualized the coding scheme, decomposing these patterns into their underlying dimensions, and analyzing each student's exploration with respect to these dimensions. We focused on the sequence settings, temporarily ignoring some of the other exploration features: action rate and observation times (will be re-incorporated in research under way). With a larger data-set, we found that only half of the students employed "pure" patterns, a shift from our previous results.

We observed a dominance of an exploration pattern described as "oscillating between extremes". In this style of exploration, the students move back and forth between polar values in the parameter space. In previous work (Levy & Wilensky, 2005) we had observed that when the exploration tools are stronger (using the command line to program rules and settings into the model), less limited by the provided widgets, this pattern became even more prominent, and was taken up by most of the students. This most common strategy provokes interesting questions. In moving between extremes, it seems that a continual comparison is made between "now" and "previous". As a model's settings change, the previous model behavior soon disappears and leaves no trace or record. If one were to search for a relationship between such changes and the resultant model behavior, pair-wise comparison between the current and the last setting emerge into an oscillating pattern, which is guided by the edges of the parameter space. A similar strategy is recommended when exploring simulations (Gilbert & Troitzch, 2005), as it allows testing the model in distinct and extreme situations.

Among the other patterns, some did not fall into a defined pattern, resulting mainly from the students failing to reach the region, where fruitful information could assist them in solving the problem.

A majority of the rest (a quarter of the students), employed a two-phase exploration. The first phase was devoted to "answering the question", an externally driven task. For this part, they employed the "direct" exploration style. Once solving the question, they changed tack and in the second phase, they devised their own experiment, an internally driven exploration, employing a different style. Thus, one may view these students as operating in different modes in guided and unguided situations.

A case study with one student: harnessing logging to gain a deeper view into learning

In the case study described, we have demonstrated how a combination of analyzing the student's actions with the models and her text-based responses helped detect significant times in the process of learning.

Using the pre-test and post-test results we detected what she knew walking in and what she learned through the curriculum. We also exposed her difficulty in connecting among levels, even in the post-test. By analyzing her model explorations we detected two sensitive times, when she changed her style of exploration. Observing her text-based answers afforded a window into her voiced and later resolved confusions.

This student, usually compliant and adapting to external direction, changed her style of exploration with the models, once struggling with the concepts at hand. Her style changed from "straight to the point" and "constant intervals" canonical patterns to "inching through space" once her prior knowledge and the external directives posed a puzzle with respect to the model's behaviors. With this style of exploration, she quickly and richly scanned the parameter space in different directions, finally converging to a region that helped her resolve the puzzle. In this region, she explored several settings.

When presented with another puzzle, she voices her confusion by answering with questions. Once again, her style of exploration shifts to "inching through space" and intense sampling of the parameter space along 20 different settings. She resolves the puzzle upon culminating this exploration with an understanding of the distinct nature of the micro- and macro-levels in the system.

This understanding is not exemplified in the post-test, as she continues to struggle with such issues. We have observed the budding but not the flowering of this new understanding.

Learning gains and curricular goals

The Connected Chemistry curriculum takes place during seven 30-45 minute activities. As such, it focuses on particular goals, yet not others. These goals are described in the introduction. Investigating the learning gains through pre-tests and post-tests afforded investigating the success of two of these goals: (a) Thinking "from the molecule up" by focusing on micro-to-macro descriptions, transitions and connections, set within a complexity perspective; (b) Mathematical modeling, interpreting graphs of a model's behavior and deriving equations from data obtained through the NetLogo model explorations.

Significant gains were made regarding both curricular goals.

The largest gains were in understanding chemical phenomena at the micro-level and their connection to the macro-level phenomena. As traditional curriculum approaches mainly the macro-level view of these phenomena, we had approached the topic with an alternative perspective, highlighting the molecules' behaviors and the causal relation between these behaviors and the classical macro-view. Regarding both goals, we have seen and important change in the students' understanding of the topic at hand.

With respect to mathematical modeling, we have seen a strong improvement in two skills: graph interpretation and constructing a relational equation from scatter-plot data. These two skills were intensively practiced throughout the curriculum. The models usually incorporated time plots, enabling a focused view of the evolving system behaviors and their relation to the particles' behaviors. Graph interpretation was necessary to approach a large portion of the targeted tasks. In the second part of the curriculum, the students were engaged in deriving the quantitative Gas Laws, by designing experiments, collecting data, obtaining scatter-plots and deriving the symbolic equations. This took place four times and is related to the students' observed significant gains in these abilities.

Finally, regarding numerical problem-solving we have seen a non-significant improvement. This is the traditional type of tasks presented for this topic in high-school Chemistry classrooms. While this is a dimension to be improved in future versions, we had not viewed our curricular package as particularly suited for this type of learning. We believe that combining model-based learning with numerical problem-solving would remedy this more successfully.

Future work

We are currently engaged in the final stages of automating most of the analysis. We intend to re-combine the dimensional analysis of students' model exploration patterns with action-rates and observation times to complete the cognitive picture of students' typical styles of approaching computerized models. These patterns will be related to overall learning gains as well as their learning during the activities.

REFERENCES

Ayala, C. C., Shavelson, R. J., & Yin, Y. (2002). Reasoning dimensions underlying science achievement: The case of performance assessment. *Educational Assessment, 8*(2), 101-121.

Bowen, C.W., & Bunce, D.M. (1997). Testing for Conceptual Understanding in General Chemistry. *The Chemical Educator*, 2(2), 1-17.

Bruner, J.S. (1973). Organization of early skilled action. *Child Development*, 44, 1-11.

Celko, J. (1999). SQL for Smarties: Advanced SQL Programming. 2nd edition. Morgan-Kaufmann.

de Berg, K.C., Treagust, D.F. (1993). The presentation of gas properties in chemistry textbooks and as reported by science teachers. *Journal of Research in Science Teaching*, 30(8), 871-882.

Eysenck, M.W., Keane, M.T. (1990). *Cognitive Psychology, a student's handbook*. Lawrence Erlbaum Assoc. Publishers: Hillsdale, NJ.

Fischer, K.W. (1980). A theory of cognitive development: The control and construction of hierarchies of skills. *Psychological Review*, 87(6), 477-531.

Frederiksen, J.R., White, B.Y. & Gutwill, J. (1999). Dynamic mental models in learning sciences: The importance of constructing derivatiaonal linkages among models. *Journal of Research in Science Teaching*, 36(7), 806-836.

Frese, M. & Sabini, J. (1985). Action theory: and introduction in Frese, M. & Sabini, J. (Eds.) *Goal Directed Behavior: The Concept of Action in Psychology*. Hillsdale, New-Jersey: Lawrence Erlbaum Ass., Pub.

Gibson, E.J. (1969). *Trends in perceptual development. Principles of Perceptual Development*. Chapter 20, 450-472. Englewood Cliffs, NJ: Prentice-Hall Inc.

Gibson, E.J. (1988). Exploratory behavior in the development of perceiving, acting and the acquiring of knowledge. *Annual Review of Psychology*, 39, 1-41.

Gilbert, N., & Troitzsch (2005). *Simulation for the social scientist*, Chapter 2. Second edition. Open University Press.

Gobert, J., Horwitz, P., Tinker, B., Buckley, B., Wilensky, U., Levy, S., and Dede, C. (2003). Modeling Across the Curriculum: Scaling up Modeling Using Technology. In the Proceedings of the Twenty-fifth Annual Meeting of the Cognitive Science Society, July 31-August 2, Boston, MA.

Goldstone, R. (1998). Perceptual learning. *Annual Review of Psychology*, 49, 585-612.

Greeno, J.G. (1978). Natures of problem-solving abilities. In Estes, W.K. (Ed.), *Handbook of Learning and Cognitive Processes*, Vol. 5: Human Information Processing. New-Jersey: Lawrence Erlbaum Associated.

Han, J. & Kamber, M. (2006). Data Mining: Concepts and Techniques. 2nd edition. Morgan-Kaufmann.

Holland, J. (1995). *Hidden Order: How Adaptation Builds Complexity*. Helix Books/Addison-Wesley, Reading, MA.

Horwitz, P., & Burke, E. J. (2002). Technological advances in the development of the hypermodel. Paper presented at the American Educational Research Association, New Orleans.

Kauffman, S. (1995). *At home in the Universe: The Search for the Laws of Self-Organization and Complexity*. Oxford University Press, Oxford.

Kimball, R., & Ross, M. (2002). The Data Warehouse Toolkit: The Complete Guide to Dimensional Modeling. 2nd edition. Wiley.

Kozma, R., Russell, J., Jones, T., Marx, N., & Davis, J. (1996). The use of multiple, linked representations to facilitate science understanding. In Vosniadou, S., Glaser, R., DeCorte, E., and Mandl, H. (Eds), *International Perspectives on the Psychological Foundations of Technology-Based Learning Environments*. Erlbaum, Hillsdale, NJ, 41-60.

Levy, S.T., & Wilensky, U. (2005). Students' patterns in exploring NetLogo™ models, embedded in the Connected Chemistry Environment. In J. Gobert (Chair) and J. Pellegrino (Discussant), Logging Students' Learning in Complex Domains: Empirical Considerations and Technological Solutions. Paper presented at the annual meeting of the American Educational Research Association, Montreal, Canada, April 11-15, 2005.

Levy, S.T., & Wilensky, U. (2006a). Emerging knowledge through an emergent perspective: High-school students' inquiry, exploration and learning in Connected Chemistry. Paper presented at the annual meeting of the American Educational Research Association, San-Francisco, April 7-11, 2006.

Levy, S.T., & Wilensky, U. (2006b). Gas Laws and beyond: Strategies in Exploring Models of the Dynamics of Change in the Gaseous State. In B. Buckley (Organizer), J. Gobert (Presider), & J. Kraicik, (Discussant), "Supporting science learning and science education reform with information technologies". Paper presented at the National Association for Research in Science Teaching, San Francisco, CA.

Levy, S.T., Novak, M., & Wilensky, U. (2006). Students' forages through the complexities of the particulate world: Scaffolding for independent inquiry in the Connected Chemistry (MAC) curriculum. Paper presented at the annual meeting of the American Educational Research Association, San-Francisco, April 7-11, 2006.

Levy, S.T., Novak, M., Wilensky, U. (2005). Connected Chemistry Curriculum 1.3 Evanston, IL. Center for Connected Learning and Computer Based Modeling, Northwestern University. ccl.northwestern.edu /curriculum/chemistry/.

Lin, H-S, Cheng, H-J (2000). The assessment of students and teachers' understanding of gas laws. *Journal of Chemical Education*, 77(2), 235-238.

Marbach-Ad, G., Stavy, R. (2000). Students' cellular and molecular explanations of genetic phenomena. *Journal of Biological Education*, 34(4), 200-205.

Mas, C.J.F. & Perez, J.H. (1987). Parallels between adolescents' conceptions of gases and the history of chemistry. *Journal of Chemical Education*, 64(7), 616-618.

McCullough, Malcolm (1996). *Abstracting Craft*. Cambridge, Massachusetts: The MIT Press.

Niaz, M., Robinson, W.R. (1992). From 'algorithmic mode' to 'conceptual gestalt' in understanding the behavior of gases: An epistemological perspective. *Research in Science and Technological Education*, 10(1), 53-65.

Piaget, J. (1972). The role of action in the development of thinking in Overton, W.F. & McCarthy Gallagher, J. (Eds.), *Knowledge and Development*, Volume 1: Advances in Research and Theory (pp 17-42). Plenum Press.

Resnick, M. & Wilensky, U. (1993). Beyond the Deterministic, Centralized Mindsets: New Thinking for New Sciences, American Educational Research Association, Atlanta, Ga.

Rowe, P.G. (1987). *Design Thinking*. Cambridge, MA: The MIT Press.

Samuelson, L.K., Smith, L.B. (2000). Grounding development in cognitive processes. *Child Development,* 71(1), 98-106.

Searle, J.R. (1981). The Intentionality of Intention and Action. In D.A. Norman (Ed.), *Perspectives on Cognitive Science* (pp 207-230). Norwood, NJ: Ablex Publishing, Hillsdale, NJ: Lawrence Erlbaum Assoc.

Shavelson, R. J., Li, M., Ruiz-Primo, M. A., & Ayala, C. C. (2002). *Evaluating new approaches to assessing learning.* Paper presented at the Keynote Address: Joint Northumbria/EARLI Assessment Conference, University of Northumbria at Newcastle, Longhirst Campus, UK.

Simon, H. A. (1978). Information-processing theory of human problem solving. In W. K. Estes (Ed.) *Handbook of learning and cognitive processes*: Vol. 5. Human information processing (pp. 271-295). Hillsdale, NJ: Lawrence Erlbaum.

Snodgrass, R. T. (2000). Developing Time-Oriented Database Applications in SQL. San Francisco: Morgan Kaufmann.

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

Thompson, K. (1968). Programming Techniques: Regular expression search algorithm

Uzgiris, I.C. (1983). Organization of Sensorimotor Intelligence, Chapter 5 in Lewis, M. (Ed.), *Origins of Intelligence, Infancy and Early Childhood* (pp 135-190). New-York and London: Plenum Press.

von Hofsten, C. (1993). Prospective control: A basic aspect of action development. *Human Development*, 36, 253-270.

Werner, H. (1957). The concept of development from a comparative and organismic point of view, pp 125-140 in Harris, D.B. (Ed.) *The Concept of Development*. Minneapolis: University of Minnesota Press.

Wilensky, U. & Resnick, M. (1999). Thinking in Levels: A Dynamic Systems Perspective to Making Sense of the World. *Journal of Science Education and Technology*, 8(1).

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

Wilensky, U. (1999b). GasLab: an Extensible Modeling Toolkit for Exploring Micro- and Macro- Views of Gases. In Roberts, N. ,Feurzeig, W. & Hunter, B. (Eds.) *Computer Modeling and Simulation in Science Education*. Berlin: Springer Verlag.

Wilensky, U. (2000). GasLab curriculum, Center for Connected Learning and Computer-Based Modeling, Northwestern University. Evanston, IL. http://ccl.northwestern.edu/curriculum/gaslab/.

Wilensky, U. (2003). Statistical mechanics for secondary school: The GasLab modeling toolkit. *International Journal of Computers for Mathematical Learning*, 8(1), 1-41 (special issue on agent-based modeling).

Wilensky, U., Hazzard, E. & Froemke, R. (1999). An Extensible Modeling Toolkit for Exploring Statistical Mechanics. Proceedings of the Seventh European Logo Conference - EUROLOGO '99, Sofia, Bulgaria.