Emerging knowledge through an emergent perspective: High-school students' inquiry, exploration and learning in Connected Chemistry

Sharona T. Levy & Uri Wilensky

stlevy@construct.haifa.ac.il uri@northwestern.edu

Center for Connected Learning and Computer-Based Modeling University of Haifa, Israel 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 2006 annual meeting of the American Educational Research Association, San Francisco, USA.

ABSTRACT

We investigate students' strategies in exploring computer models and the way these relate to the underlying mathematical behavior of the model, and to prior knowledge in the domain.

In the Modeling Across the Curriculum project (Gobert et al, 2003), we enable students' exploration of computer models that are embedded in a supporting script. The Connected Chemistry environment (Levy, Novak & Wilensky, 2005) is one such model-based curriculum. Connected Chemistry (CC) employs multiagent NetLogo models to enable students' inquiry, manipulation, and observation of chemical "entities" at the molecular level and the resulting group-level or "aggregate" patterns. The first CC chapter is on the topic of Gas Laws and Kinetic Molecular Theory. In the supporting script, we facilitate students' inquiry into these topics through a more general complex systems lens.

We have studied students' strategies when exploring three models, with the goal of extracting a quantitative relationship among the variables. These models differ in their mathematical behavior (linear versus inverse functions). In addition, we have tested the association between these strategies and the students' prior knowledge, as evidenced in the pre-test results.

We have found the following results: (1) The students used mainly a "constant intervals" strategy when exploring one of the models, which is based on a linear relationship between the variables (number and pressure); For the model based on an inverse function (volume and pressure), they used mainly an "increasing intervals" strategy; Curiously, they used an "increasing intervals" strategy when exploring the second model, which is based on a linear relationship (temperature and pressure). (2) Students were consistent in the strategy they employed when moving from the linear temperature & pressure model to the inverse volume & pressure model. (3) Prior knowledge is associated with the strategy the students use in exploring the three models.

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, 1990; 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 & Wilensky, 2004; Levy, Novak & Wilensky, 2005; 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 TMscript (Horwitz, 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.

In the current study, we focus on three activities: Number and pressure, Temperature and pressure, Volume and pressure. Three explorations were selected, which share a common structure; the students are extracting data from the models in order to construct the Gas Laws equations. They differ in the underlying mathematical behavior: two explorations involve a linear relationship between the two variables (number of particles and pressure, temperature and pressure), while the third involves an inverse relationship (Boyle's law - volume and pressure). For a linear function, it makes sense to use a commonly taught strategy of even additions to the independent variable. However, when exploring an inverse function, this strategy is less fruitful. Since the function changes quickly for low values of the independent variable and slows down as larger values are reached, a strategy that results in dense data points in the low range would provide more information regarding such The exploration strategies are accessed via the log data, using automated algorithms, which detect the sequence of settings. Students' prior knowledge is obtained from the pretest results.

In previous work (Levy & Wilensky, 2005), we have detected and described four types of sequences, which are employed by students: constant intervals, increasing intervals, decreasing intervals and mixed sequences. In the current study, we continue this line of research one more stage, by using a much larger sample and observe whether these strategies are associated with prior knowledge.

In all three activities, after qualitative exploration of the models takes place, the students set out to derive an equation relating the aggregate gas variables, the Gas Laws. They collect five data-pairs from the model, regarding the two variables and place them in a table. Following this, a scatter-plot of the data-pairs is displayed, the students describe the relationship qualitatively, and then in the form of the appropriate equation. This equation is then used in subsequent tasks to solve for typical chemistry problems.

The following screens (Figure 1) demonstrate the focal screens in two of the explorations:

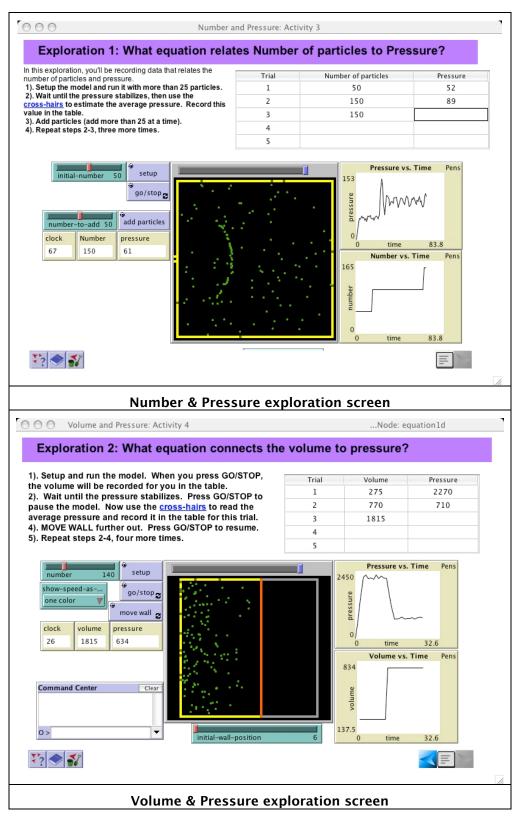


Figure 1: "Number and pressure", "Volume and pressure" quantitative exploration screens.

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. In the future, will report on their relation with the text-based answers, reflecting prior knowledge, knowledge-inconstruction and learning outcomes.

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 2) 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:

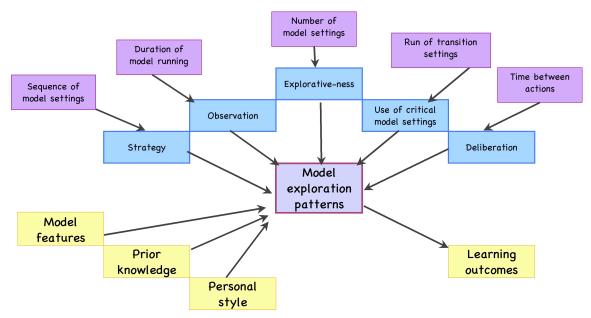


Figure 2: Framework for analyzing students' activities with the Connected Chemistry models, and their relationship to learning outcomes.

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.

In this study, we focus only upon the students' strategies in selecting the settings with which to run the model.

Focus

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

- (a) Characterizing students' model exploration when the goal is to extract data to construct an equation.
- (b) Comparing students' exploration of the model's parameter space, when different mathematical relationships underlie the target system's behavior.
- (c) Exploring whether students' exploration strategies may be associated with prior knowledge.

METHOD

Sample

Our sample includes 480 high-school students from two schools. 143 students come from one school, which is highly diverse with over 20 spoken languages and 20% free lunch. 337 students come from another school, with similar diversity – 16 spoken languages and 6% free lunch.

Data collection

The schools access the activities via the web and the students' interactions are logged from afar. We collect students' work with the CC scripts and models through logs of their text-based activities, as well as their actions in manipulating the models. As the activities progress, the students gain fluency with the investigation tools and the inquiry scaffolding is faded.

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 this paper, we focus on the latter – what characterizes the students' exploration of the Connected Chemistry models, themselves.

The students filled out a pre-test and post-test questionnaire which addressed several dimensions targeted in the curriculum: understanding the microscopic gas particle behaviors, the aggregate systems behaviors in qualitative and quantitative terms, micro-to-macro descriptions of the system, the dynamics of change, randomness of gas particle behavior, graph interpretation and deriving an equation from a graph.

Analysis

The students' entries into the three tables (see Figure 1) were collected from the logs. The first and second derivatives of these sequences were calculated. The second derivative was then recoded as its sign – positive (indicating increasing intervals), negative (indicating decreasing intervals) or zero (constant intervals). For the Volume & Pressure model, in which the students click on the model itself to determine the volume, up to a 5% from linearity was calculated as linear (based on the mean of the first derivative). The 27 combinations of the three signs of the second derivative, were sorted into four categories: mainly constant intervals (at least 2/3 constant additions), mainly increasing intervals (at least 2/3 decreasing additions), mainly decreasing intervals (at least 2/3 decreasing additions) and mixed (when no trend is detected).

The students' pre-test score is the number of correct answers (out of 18).

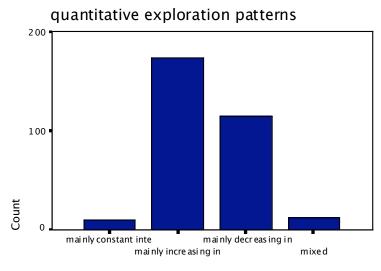
RESULTS

Patterns in exploring models

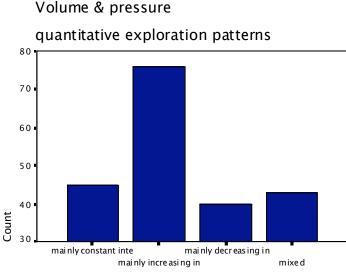
We present the results in the following Figure 3. Statistical analysis is still underway. We can see that the dominant strategy for the linear Number & Pressure exploration is that of constant intervals. For the linear Temperature & Pressure exploration, there are two dominant strategies: Increasing intervals and Decreasing intervals. Finally, for the inverse Volume & Pressure exploration, the dominant strategy is that of Increasing intervals. Thus, we conclude that different mathematical behaviors elicit different model exploration patterns. However, given the results for the Temperature & Volume model, these are not associated solely with the mathematical behavior of the model. Additional factors impact the students' search for information. Number & pressure quantitative exploration patterns

NP sma**l** group

Temperature & pressure



TP small group



VP small group

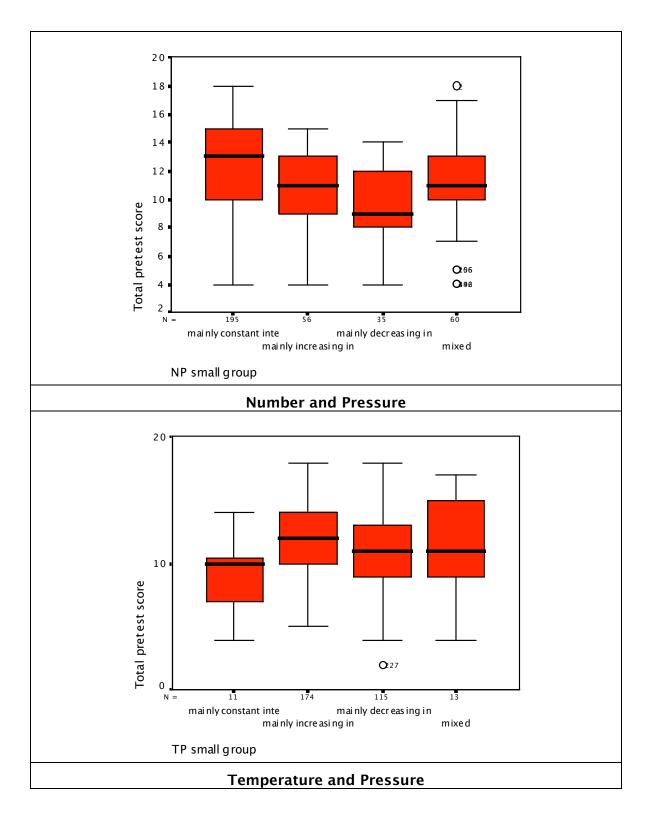
Figure 3: Distribution of strategies in exploring Connected Chemistry models, which are geared towards quantifying a relationship

Consistency in the strategies

We have seen a departure from the dominant "Constant intervals" strategy demonstrated in the Number & Pressure activity when approaching the following two activities. We have found that while there is a larger spread among the strategies, there is a strong consistency within each student. Cramers' V is 0.191 with a significance of 0.017, and Phi is 0.330 with the same significance.

Prior knowledge and exploration strategies

The following Figure 4 illustrates the association between the different patterns and the students' pre-test score. We can see that in the Number and Pressure activity, the dominant "Constant intervals" strategy is associated with higher pre-test scores. For the other two activities, the dominant "Increasing intervals" strategy is associated with higher pre-test scores.



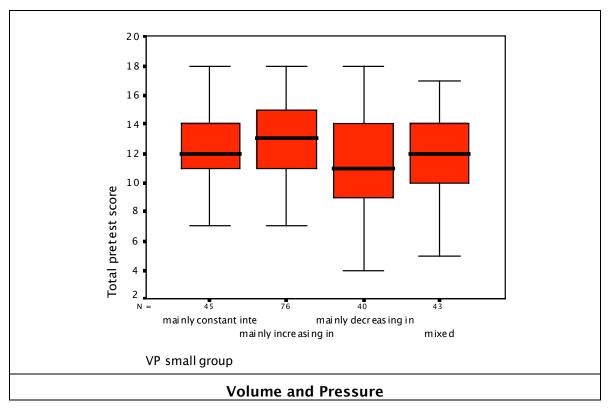


Figure 4: Association between exploration patterns and pre-test scores

To summarize this study, we have seen the following:

- (1) In an activity aimed at deriving a quantitative relationship, we have seen the use of two dominant strategies: "constant intervals", where the independent variable in the experiment is increased at constant intervals; and "increasing intervals", where the independent variables is increased at increasing intervals.
- (2) In one of the models with a linear function underlying the model's behavior (Number and Pressure), most of the students used the "constant intervals" strategy.
- (3) In a model with an inverse function underlying the model's behavior (Boyle's Law), almost half of the students adapted their exploration to the inverse function and a wider distribution of strategies was observed.
- (4) In one of the models with a linear function underpinning the model's behavior (Temperature and Pressure), most of the students employed an "increasing intervals" strategy. The students were consistent in their search with the strategy they employed to derive the inverse relationship.
- (5) The use of the different strategies is associated with different pre-test scores.

CONCLUSION

How do students search for information within computer models? We have found that underlying model behaviors as well as additional factors interact in shaping the particular form by which information is searched for. These forms are described via a multi-faceted framework (Figure 2), 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 models in which they collect data aimed at deriving an equation.

The "Constant intervals" pattern was observed when the students were recording data in a table to be used in the next screen in order to derive an equation. It is described as a constant change to either the independent or dependent variable in the experiment, e.g. 10, 15, 20, 25... A linear addition to the independent variable is the most commonly taught strategy in science classes. Constant additions to the dependent variable may seem quite wrong, as this runs contrary to normal science inquiry teaching. However, when the underlying functions are not linear, as in Boyle's Law, if one wishes to study the changes in the system behavior – it would make more sense to sample evenly for the dependent variable, rather than independent variable. This way, one captures the full range of change in the underlying behavior. Nevertheless, it is not only the mathematical behavior, which impacts the students' search for information. We have seen that for one of the models, which displays the linear relationship between Temperature and Pressure, the students consistently used the same strategy they used in the Volume & Pressure inverse relation model.

How does the model's underlying mathematical function affect the students' exploration patterns? Do they use the commonly taught "constant additions" for the independent variable regardless of the rate of change of the dependent variable? As described above, commonly taught practices in science inquiry do not always fit for every context. If one didn't know the relationship between the variables, linear variation of the independent variable is a good heuristic for mapping out the space. However, this is not the case in our investigation. Prior to this quantitative section, the students have interacted with the model and attended to its qualitative relationship between volume and pressure in Boyle's Law, they may notice that the pressure changes faster when the volumes are smaller. In this case, evenly spacing the values for the dependent variable, rather than the independent variable will give a better sense of the full parameter space.

We have found that more than half the students adapted their exploration to the inverse function underpinning our "Volume and Pressure" model. They either consistently increased the increment size or hybridized a linear strategy,

"breaking the slope" at some point to increase the increments. We were quite taken with the results. Inverse relationships are typically harder to learn, and we have a "linear" bias in our spontaneous view of the world (Nemirovsky, 1994). However, in this case we have seen the students internalize the inverse relationship, even without numerical and symbolic forms, simply via model manipulation and observation. This internalization is evidenced in their adaptation to the specifics of the function.

At this point, we are still in the process of analyzing and interpreting these results. We intend to break up the analysis of the pre-test scores by dimensions (e.g. are these strategies associated more strongly with the scores on problems, that involve quantitative calculations?). The post-test results enable us to ask whether these strategies are associated with learning gains.

REFERENCES

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.

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.

Gobert et al, 2003

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

Horwitz, 2002

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

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.

Kozma, Russell, J., Johnston, J. and Dershimer, C. (1990). College students' understanding of chemical equilibrium. A paper presented at the Annual Meeting of the American Educational Researcher Association, Boston, MA.

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/.

Levy & Wilensky, 2005

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.

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.

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

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

Uzgiris, I.C. (1983). Organization of Sensorimotor Intelligence, Chapter 5 in Lewis, M.

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