Part II: FINDINGS

I. Overview of Findings

> *"Oh, wow, I see how it works. So, like the air molecules are hitting our skin right now?"*
>
> *- Biology student*

The overall goal of the research was to investigate the extent to which students can relate their understanding of the interactions of atoms and molecules, usually the domain of physics and chemistry, to interpretations of biological phenomena through the use of molecular dynamic models. An additional goal was to understand how teachers implemented model-based materials in their classrooms.

Biology, now more than ever, employs chemical and physical principles for explanations of phenomena at the molecular level. As suggested at a meeting of the advisory board, certain physical-chemical principles were selected that underlie many biological processes. A deep understanding of these ideas would constitute a form of “Molecular Literacy” needed by any student to understand the molecular logic of biology. These activities were labeled "Stepping Stones to Molecular Literacy" and they became the focus for project assessment and evaluation.

The project objectives for the first year were to develop Stepping Stone activities and the capacity of the *Molecular Workbench*. The activities were to be linked to both standards and typical textbooks, able to stand alone, and easy to implement in a variety of styles and in diverse educational settings. In addition, the assessments and their delivery system (the database) were all started during the first year and a half of the project that began in February, 2003.

Formative testing of the new materials occurred during the 2004-05 school year. The formative test teachers were ~~core~~ high school biology teachers. These teachers inserted the Stepping Stone activities, and some others, during the school year. The teachers administered a pre- and post-test and used the activities at this point in the project at times of their choosing.

The summative tests, composed of a “Short Sequence” of activities, took place during the Fall of 2005. The activities also had a sequential molecular logic. In addition to the short study, a controlled study of an activity developed with and without a dynamic model was completed.

II. Classroom Test Sites

Over the two years of formative and summative research, 25 teachers tested activities in 74 classrooms (19 urban classrooms, 37 suburban and 18 rural classrooms). Approximately 1300 students used the materials as part of the various studies described below.

Each stepping stone activity took between one and two 50-minute class sessions. The pre-/post-test took one class session each. On average, students experienced Molecular Logic (MoLo) activities for 5-6 days per semester. In addition, the teacher feedback forms averaged 15 minutes for completion per activity and teachers were expected to participate in a 30-45 minute telephone interview.

Teacher Recruitment and School Characteristics

Teachers were recruited through the AP biology list-serve, conferences, and the Concord Consortium website. Teachers with prior working relationships with the Concord Consortium were also recruited. A range of types of classrooms guided the selection, including demographic distribution (urban, rural, and suburban), national distribution, as well as level (AP, honors, general biology). In addition, the schools had to have minimum technical requirements of Windows 98 or later or Macintosh OS9 or later as well as an on-site network server connected to the Internet. The selected schools were distributed nationally, located in Arizona, California, Colorado, Connecticut, Kansas, Maryland, Massachusetts, Michigan, Minnesota, Mississippi, Missouri, Montana, North Carolina, and Wisconsin.

For the 2004-05 formative field test we recruited 22 teachers to participate in the formative testing. **Core field-test teachers** were asked to test the entire set of 10+ "Stepping Stone" activities in the database, administer pre- and post-tests, ~~and~~ provide online feedback about the activities in the class and participated in telephone interviews. A **free-choice field-test** teacher was asked to search the database and select five model-based activities of their choosing, test them in the classroom, and give us feedback, including a phone interview. Core test teachers enabled us to review the logic of the stepping stone activities and their value, whereas the free-choice teachers were useful in highlighting the activities that teachers would choose on their own.

We tested the materials in 48 classrooms, totaling approximately 900 students using the materials as part of the core field test study. Classrooms included 35 tenth grade classes, five eleventh grade classes, two twelfth grade classes and two mixed grade classes. They were classified as:

- 14 Urban classes with a total of 206 students
- 22 suburban classes with a total of 495 students
- 12 rural classes with a total of 199 students

For the 2005-06 summative test, nine teachers, three of them new, tested the materials in 26 classrooms, totaling approximately 400 students using the materials as part of the implementation study. They divided into:

- 5 Urban classes
- 15 suburban classes
- 6 rural classes

There were 13 tenth grade classes, seven eleventh and twelfth grade classes (some were mixed), three ninth grade classes and one tenth-twelfth mixed grade class.
Teachers were offered stipends for field-testing the activities, administering a pre- and post- test, participating in a half-hour long telephone interview each semester and for providing written feedback for each activity using an online feedback form.

This summative test was limited to one semester, and was composed of a subset of the molecular Stepping Stone activities that had not been extensively tested before or had undergone a radical revision as a result of the formative tests.

Teacher Characteristics

All selected teachers completed an initial survey on their educational background and their attitudes towards science and technology.

Biology teachers were our target teachers; all test teachers had biology undergraduate degrees. Additionally, 60% of the teachers had advanced degrees in biology and/or science education. The teachers were all very comfortable using computers. It should be noted that on a Likert scale of one to five where five indicated full agreement, a value of 4.2 was obtained for teachers' confidence in using additional technology in the class.

The teachers who participated were self-selected. The biology teachers clearly had strong connection to the content, (we did not have teachers teaching outside of their subject area), and were interested in new methods for helping their students learn the material. Based on the above analysis of the teacher background surveys, these teachers were confident both with the subject and with incorporating technology in their classrooms. Though these teachers do not necessarily represent the full range of biology teachers, the field test teachers did represent the range of classrooms within which the materials could be used.

Course Characteristics and Teacher Ease

The initial survey indicated that the teachers were quite comfortable with teaching topics in molecular biology (mean value of 4.26 on a Likert scale), less so with chemistry (mean value of 3.73), and even less so in physics (mean value of 2.73). This was expected. MoLo was offered in 74 biology courses with 11 different titles. About 30% of the courses can be considered to be lower level Biology 2 and General Biology courses, while the remaining 70% can be considered to be middle or upper level courses such as College Prep Biology, Honors Biology, and Advanced Placement Biology. It appears that teachers expected that our material was a good fit to the emphasis on molecular biology in the upper level classes.

Teacher Compliance and Feedback.

The 2004-05 field test teachers tended to select among the activities rather than use them all. Twenty-seven percent of the teachers did between 8-16 of the Stepping Stone activities (Note some of the Stepping Stones contained several activities); 45% of teachers did 5-7 of the activities; and 28% of the teachers completed 2-4 activities.

Each teacher was asked to complete and submit an online feedback form after they used a Stepping Stone activity in their classes. The online feedback forms were set up for teachers to provide feedback on the activities they were teaching and how the activities were used in the classroom. Teachers were also asked to inform us about any technical difficulties they experienced and to record typos, awkward phrasing, and passages of text that were difficult for their students. Teachers were asked to complete these forms after completing each activity.

Nineteen teachers submitted a total of 92 online feedback forms (a handful of the feedback forms addressed multiple activities). Though the teachers made an effort to complete the forms, they did not complete them for all the classroom activities they used. Feedback was given on 96 activities, but we know, based on submitted student reports, that at least 139 activities were completed. Thus, feedback was obtained from teachers on no more than 69% of the activities used.

For the Fall 2005 Short Sequence Study, the nine test teachers completed all they were asked to do—four sequentially ordered and revised Stepping Stone activities. Again teachers were asked to give feedback on a revised online feedback form. We received 32 feedback forms, an 88% compliance rate.

Data Sources

1. On-line feedback forms. A website was set up with a questionnaire for each teacher to complete after implementing an activity with their classes (see *Appendix F: On-line Feedback Form)*. We received reports not only on the Stepping Stones, but also on several other activities. The online feedback forms were analyzed for insight into classroom implementation issues, technical difficulties, and suggestions for activity revisions.

2. Telephone interviews were held with teachers (See *Appendix G: MoLo Telephone Questionnaire ~~Interview~~*). An interview protocol was used to guide the phone calls, but conversations went beyond the limit of the protocol when greater detail about teachers' thoughts on the activities was warranted. Responses to the telephone interviews were used to advance our understanding of classroom implementation and user interaction of the database and the models. The telephone interviews were conducted once for each teacher at the end of the field study. The telephone interviews contained 16 questions and lasted from 30-60 minutes.

3. Classroom observations were conducted in local Molecular Logic ~~(MoLo)~~ test classrooms. MoLo team members attended classroom implementation of Stepping Stone activities specifically to observe user interaction with the model-based activities. Where did students get confused? What technical difficulties arose? Did students complete the activities? There were five teachers visited in Massachusetts in the first year and one in the second year. A minimum of two classroom visits was done with each local teacher. In one case five classroom visits were completed. Notes from these observations were distributed to the project staff.

4. Student work. Each model-based activity produced reports that were either uploaded directly to the Concord Consortium server or were printed by the students and mailed to the researchers at Concord Consortium. We received 2531 uploaded student reports and ~1850 hard copy student reports (See *Appendix H: Sample Report*). The reports included students responses to embedded multiple choice and open-ended assessments as well as snapshots of models that the student recorded when a specific phenomenon occurred. The snapshots are valuable tools used to give researchers and teachers insight into what students are able to recognize.

5. Student feedback. We conducted a classroom discussion at the end of the year with one local class and we received written feedback from students in a several classes.

Activity Review and Revision

Activity review reports were written for each Stepping Stone activity. The reports drew on several sources of data: student activity reports, teacher feedback forms, and classroom observations. Student activity reports were the largest source of data for each activity. Our online database of uploaded reports made it easy to take a semi-random sample of reports and analyze them. Reports representing the cross-section of levels, geographic distribution and demographic distribution were sampled. The analysis looked at students' responses to multiple choice and open-ended questions as well as snapshots of the models and their annotations. For each question, student answers were grouped into categories and tallied. This allowed us to diagnose which questions were most challenging, and specifically what confusions were arising.

Included with the reports were action points that focused on revisions that needed to be made immediately, such as rewording of text and questions, technical errors and bugs, plus long-term recommendations for restructuring portions of activities and models to better meet the objectives of the activities. Suggestions that could be repaired easily were addressed immediately. The longer-term suggestions were compiled and activities were further revised.. Because the activities are distributed from a central database, corrections and other changes can be promulgated immediately to all participating teachers.

The feedback forms which teachers completed after each activity often included comments on specific models, questions, or vocabulary within an activity. We increased the specificity of teacher responses part way through the project by changing the survey to include a section asking them to select and comment on a single page of the activity.

Classroom observations provided a crucial source of data not available through other means. Watching students work with the software often revealed surprising behaviors and unexpected confusions.

Based upon the results derived from these various techniques, changes were made to the database, the Stepping Stone activities and to the *Molecular Workbench* software itself.

III. Classroom Implementation

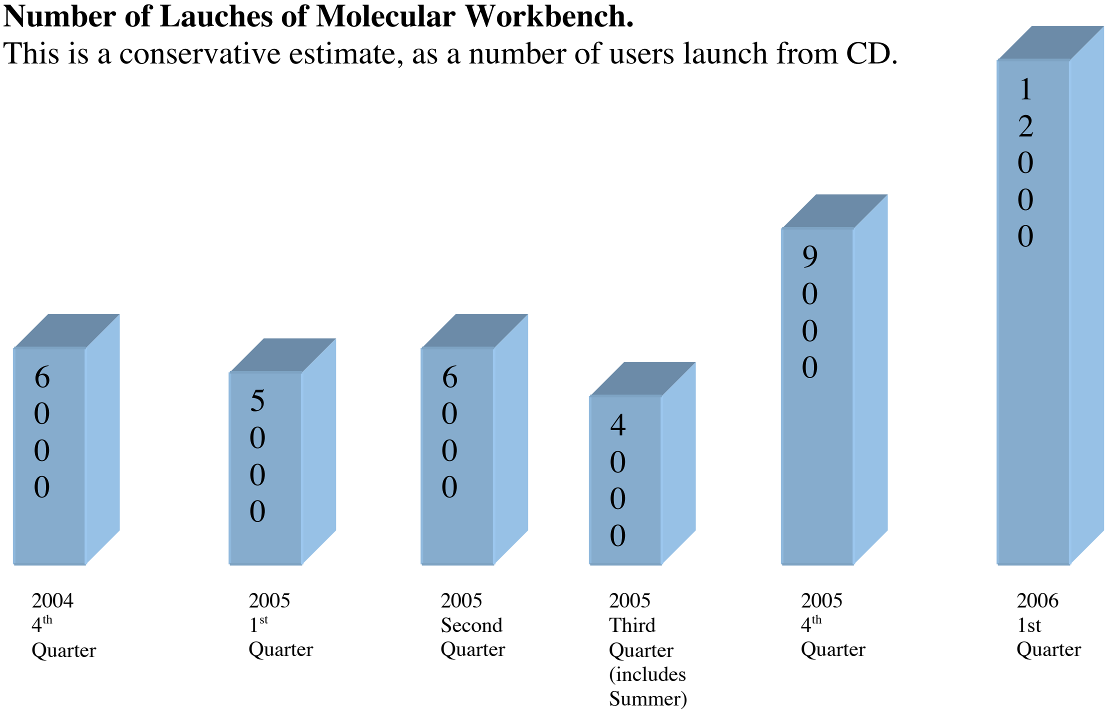
Technologies

At the project start there were technical challenges in activity delivery, including getting past school firewalls and writing to school computers. A variety of solutions were developed and made available in the FAQ area of the activity database. Downloading *Molecular Workbench* onto computers and then working off-line solved the remainder of the technical difficulties for teachers.

By the Fall of 2005, 72% of the schools were having no technical difficulties; the remainder had problems mostly unconnected with *Molecular Workbench*. Problems that remained that last "semester" included a school whose computers were infected with a virus that caused our servers to block their access, and one school that had disabled Java. Tech support at the latter was able to use the detailed installation instructions posted on the website to turn this feature back on. One bothersome problem, the running of activities that included Flash, developed at the end of the project, and affected several Stepping Stones. We are promised this problem will disappear with the next update of Java Webstart.

Number of Lauches of Molecular Workbench.

This is a conservative estimate, as a number of users launch from CD.



Use of *Molecular Workbench* increased over the school year (it tended to fall off in the summer.) There was one other project using the materials intensively, and a number of users around the world accessing *MW*. Although we cannot yet access numbers for individual activities (we will soon be able to do so), we believe much of the gain is from Molecular Logic users.

Classroom Use

The activities in the Fall 2005-05 field test were used in one of the following four ways:

> Demonstration/discussion - The teacher controls work in the classroom, projects the activity, sometimes using a Smartboard, to work through the activity ~~of the~~ as a class. This occurred about 7% of the time.
>
> Teams on Computers - Students work in pairs on computers to complete the activity. This occurs approximately 31% of the time.
>
> Library (Computer Lab) -Students work by themselves on computers after a teacher introduces the materials. This occurs about 27% of the time.

Homework (or on own time in the computer lab) - The teacher introduces content and has students work on the activity at home or in computer lab on their own time. This occurred about 33 % of the time.

Activity Selection

In the 2004-05 field test, the most commonly selected Stepping Stones activities by the biology teachers were:

- Diffusion, Osmosis, and Equilibrium – 86% participation by teachers
- Random Motion – 63%
- Tree of Life: Macromolecules – 63%
- Protein Folding – 54%
- DNA to Protein Synthesis – 59%

Teachers reported anxiety about using activities addressing unfamiliar content, such as weak forces of intermolecular attractions and protein folding. These topics not only lay outside their typical curricula, but also were unfamiliar content territory for some. In general, once teachers used them and saw their applicability, there was considerable enthusiasm for these topics.

Materials selected tended to be that which most closely matched the earliest chapters of most biology textbooks, in which chemistry and physics foundations are laid. The least selected activity was Strong Chemical Bonds, with only 13% of teachers using it. In analysis, this activity had the most technical difficulties. (The difficulty with launch problem has since been fixed). It should be noted, however, that for those activities selected, on a Likert scale of five, teachers for the most part would use the activity again (mean of 4.30).

MoLo was designed as a project that would encourage the inclusion of crucial molecular level physics, chemistry, and molecular biology concepts in the teaching of traditional Biology. We anticipated some resistance, and were not surprised when almost all of the test teachers at some point during the telephone interviews of the second year discussed how the content for many of the Stepping Stone activities lie outside the typical biology curriculum. Pressures of state standards and state tests, clear boundaries among the science disciplines, and materials going into far more detail then teachers think is appropriate for their students highlighted some of the stress teachers had in incorporating the Stepping Stones activities.

In particular chemical bonds, attractions among molecules as well as protein folding are all typically not covered in high school biology curriculum and proved to be challenges for the teachers to include. Teachers made comments such as:

> *"Strong Chemical bonds we typically don't get into that much detail, but having the visual was helpful, we talked about water and polarity and this content was reinforced. I think will help the kids understand many different*

> *topics in biology. I think it would be best if I can get the chemistry teachers to do it next year."*
>
> *"I don't usually get to protein folding. Not sure if my Biology 1 kids will be able to really understand it because we don't really do it in the class but I am going to try anyway."*

Once they used the activities, teachers did see the merits of including molecular models that illustrate the dynamic interactions of atoms and molecules and how they might support biological phenomena.

> *"We don't necessarily talk bout movement of atoms, however I can see how it helped them to understand so many thing I was teaching better. It helps them appreciate more what they are studying. The physics can only help."*
>
> *"The protein folding activity was difficult, only because I normally don't get into that detailed a look at protein. The chemistry also was hard, some of the terminology I had to take a little time to make sure I incorporated the terminology throughout my curriculum. By the end of the year though, it clearly made a difference in students thinking and communicating about content they were studying."*

More professional development around inclusion of the models and the logic behind the Stepping Stones would only improve the willingness to include these foundational models.

An Online Molecular Logic Tutorial

A Molecular Logic tutorial was created and a set of seven teachers used it and gave us feedback. It was delivered in Blackboard (See http://blackboard.concord.org, logging in as "guest" and "guest"). Its key features were assignments; a reading library; and a discussion board that included a final wrap-up survey. (See *Appendix D: The Online Tutorial*) The tutorial was revised on the basis of the feedback. Teachers were asked to help us determine what was useful and how it might work with teachers new to the MoLo activities.

Six of the seven who signed up to review this tutorial completed it and sent in their feedback. Teachers were most eager for information regarding classroom integration than for specifics in model adaptation or model building.

We anticipate continuing to develop this tool into a course introducing "Molecular Reasoning," The course would help teachers find the interconnections among the activities o that students would feel more facile moving among them.

IV. Materials Development: Lessons Learned about Modeling Layout and Pedagogy

As activities were iteratively tested, reviewed, revised, and re-tested, a number of effective pedagogical strategies were found that apply to all of them:

Making Investigatory Activities that Focus Inquiry

In designing an activity, a crucial initial step of the process is developing one or more models that create an appropriately sized investigatory space. This step sometimes has required extending the modeling engine to accommodate new topics. For activities to be successful, we developed some particular interactivity so that students can change parameters and experiment. On the other hand, reducing the number of model controllers on a page appears to improve student understanding of the concepts being modeled.

Clarifying Controls

Once the model was built, we found it important to include exactly the set of model controls that students need for that model and no more. This was particularly critical for young students. Extra controls distracted students from focusing their inquiry. Labeling each control in a key or legend and/or describing how to use it proved helpful.

Minimizing Text

Our activity pages contain a moderate amount of textual information in several categories: specifics about the model's representation, instructions for manipulating the model, explanations of the science underlying it, its direct biological relevance, and information about broader connections in biology. Activity design, therefore, was constantly a compromise between the desire to include rich support and contextual material, and the need to not overwhelm students with text as well as graphs.

The primary revision strategy has been to reduce text. We have heard from many teachers and students, and the staff observed that most students simply will not read very much on the screen, especially where there are interactive elements to manipulate, and where the text does not contain information critical to the task at hand. This reluctance may be especially strong on pages with mixed media.

Integrating Visual Elements

Activities were visually structured with bold-faced headings and sections, including clearly marked definitions, with text divided into manageable chunks. The order of elements on each page, for example, including the instructions for a model above the model on the page, was arranged so that students could look at the instructions before becoming engrossed in the model. Page titles, page numbers, and, for larger activities, a table of contents, were used to give students a sense of the larger structure while they are inside an activity, and to verbally reinforce the learning.

Designing Challenges, Hints, and Questions

Interactive model-based activities are a compromise between giving the students open-ended inquiry tasks, which can be fun and rewarding, and ensuring that they see and do tasks that are relevant to the activity. A numbered sequence of instructions that introduce the controls they'll be using, and making at least one open-ended step with a problem to solve proved a comfortable compromise. If the solution is visible in the model, students were often asked to take snapshots of their results; these are included in the report created at the end of the activity. Students were also be given space to explain how they solved the challenge.

Hint buttons help students through challenges. In several activities, hints provided several levels of guidance, starting from a more general tip to point them in the right direction, progressing to very detailed instructions for manipulating the model.

In reviewing students' answers to the embedded assessment questions, we found that they perform best on questions that are explicit, divided into simple parts, and that include specific instructions, as would be expected. More open-ended questions, on the other hand, provided staff with better information about the students' mental models, and in particular, their retention of imagery from previous models.

In general, formative testing revealed that the set of Stepping Stone activities provided fundamental concepts with unique visualizations that connect well to biological phenomena. Some improvements, however, were necessary.

V. Student Learning Gains

Formative Field Tests in Year Two

Pre-Post Learning Gains

One measure of student learning gains is the comparison of pre- and post-tests. A pre-test was administered at the beginning of the field test and the post-test was administered at the end of the field test year. The questions were designed to cover content in which students were asked to reason about the interactions of molecules that underpin the biological phenomena being modeled. Students were asked to give brief answers and to draw diagrams. The tests did not count towards student classroom grades.

All test items for modules were short-answer questions for which a rubric was created; Molecular Logic staff scored test items. Two members scored a subset of tests in order to evaluate inter-rater reliability, which was r=87% for the 2004-05 study and 88.5% for the Fall 2005 analysis.

Below are paired t-test scores for eight of the 2004-05 field study classes. Classes were chosen from those who completed a majority of the stepping stone activities. We included upper- and lower-level classes, as well as a distribution of classes from urban, suburban, and rural locations.

Year Two Pre-/Post Learning Gains Test Scores

Class	**Mean Diff**	**DF**	**t-value**	**p-value**	**Class level**	**Demographic**
1	13.471	16	11.68	<.0001	AP Bio	Suburban
2	12.342	18	10.511	<.0001	AP Bio	Suburban
3	4.667	18	4.106	.0034	AP Bio	Suburban
4	1.036	13	3.366	.0051	Standard Bio	Urban
5	9.792	11	10.48	<.0001	Honors	Suburban
6	6.5	16	6.842	<.0001	Honors	Suburban
7	6.188	15	3.807	.0017	Honors	Urban
8	3.962	12	3.749	.0028	Bio 2	rural
9	9.526	18	7.397	<.0001	AP Bio	Urban

Overall, these students show a significant increase in learning, at least to <.05 significance.

Post-test scores. For the above classes, an analysis of variance (ANOVA) was computed in order to determine whether there were any statistically significant differences on the total post-test score based on gender and on class level. The analysis revealed that the higher classes (Honors, AP) achieved predictably higher post-test scores (Mean diff 12.5, DF=116, t value=10.638, P value =<.0001) than did general students.

Gender was also significant. The analysis of data based on gender achievement showed that girls performed significantly better than boys (Mean diff 2.774, DF=116, t value=2.052, P value = .0424).

Pre/post test scores (Gain). There were, however, no differences in amount of gain, that is, the difference between pre- and post-test scores, based either on the honors/general level or on gender. That is, female students and honors students tested higher at the start and maintained their lead.

Summative Field Testing in Year Three

In the Fall 2005 field test we recruited thirteen teachers to participate in several ways: a short sequence field test or controlled test. We also asked several to review the on-line tutorial. (Two teachers dropped out of the summative test.)

The Short Sequence Field Test

The Short Sequence Field Test involved piloting four activities that made up a subset of the original Stepping Stone activities and that were logically connected. Students began the sequence by looking at macromolecules, and the weak forces that shape them. They proceeded to explore the role of different amino acids in protein-folding and continued on to explore the genetic code and how it defines the sequence of amino acids in protein.

The field test was limited to one semester because of the cycle of funding for this grant. Teachers who chose to participate were limited by the scope of the activities and how they matched the curriculum scope and sequence. Some teachers could not participate because two of the activities we needed tested covered content usually introduced in the second half of the year.

The scores below represent a random selection of the classes that completed the short sequence field test. This represents roughly half of the classes.

Year Three Pre-Post Learning Gains Test Scores

Class	**Mean Diff**	**DF**	**t-value**	**p-value**	**Class level**	**Demographic**
1	12.000	21	9.3222	<.0001	Honors	Suburban
2	8.444	17	7.696	<.0001	General	Rural
3	7.211	18	5.072	<.0001	AP-Biology	Suburban
4	5.184	18	4.120	.0006	Bio II	Suburban
5	5.441	16	3.057	.0075	Bio II	Suburban
6	10.348	22	9.268	<.0001	Honors	Suburban
7	5.708	11	4.365	.0011	Bio I	Urban
8	10.800	9	6.833	<.0001	General	Rural
9	1.265	16	2.978	.0089	General	Suburban
10	2.714	14	3.956	.0014	General	Suburban
11	9.269	12	5.285	.0002	General	Urban
12	7.167	11	4.458	.0010	Honors	Suburban

In all tests there was a significant increase in score from pre- to post-test ($p < 0.05$).

Across all students in this test there was no significant difference in the post-test scores between honors classes and general classes, unlike the difference that appeared in the year-two test. There was a significant difference between urban and rural ($p < .0001$), with the rural doing more poorly, and between suburban and rural ($p < .0001$), with rural again doing less well, but not between urban and suburban ($p = 0.15$).

The Controlled Experiment: Learning protein-folding with models vs. without

A controlled experiment was carried out to test the effects of models on the learning outcome in one stepping stone activity. In this small control study, the plan was that a biology teacher with two classes at roughly the same level would offer to one class the model-based activity (Protein-Folding) and to the other, an activity on protein folding that did not include the molecular dynamic model as its core.

The protein folding activity was chosen for the experiment because it includes visually-rich models, both a 2D protein-folding model and 3D protein structure models. The activity covers content areas that include the effect of charge on folding, the molecular basis for hydrophobicity, the effect of hydrophobicity on folding, the effect of the solvent (water vs. oil), the effect of amino acid substitutions, and the molecular basis for sickle cell anemia.

A "control" activity was constructed with the same science content, but each model was replaced by a rich set of pictures or series of pictures of states of the model. The embedded assessment questions, many of which ask students to interact with the models, were altered as little as possible. (To run both versions go to http://molo.concord.org/x)

Students in the experimental condition received the activity with models. Students in the control condition received the activity without models. Each participating teacher ran the study in at least two classes, each of which received either the experimental or control activity. Identical pre and post-tests were given, consisting of five questions (three short-essay answers and two drawings of polypeptide chains like those seen in the protein-folding model).

Three teachers participated, with 134 students in total. In all, five groups of students where both pre- and post-tests were administered (three experimental and two control groups), there was a significant increase in score from the pre- to the post-test ($p < 0.05$), but no significant difference between the experimental and control groups.

When the analysis was broken down by individual questions, across all students, the experimental group scored significantly higher on two of the five questions than the control group ($p < 0.05$). These two questions, questions four and five, were the hardest ones on the test (mean scores were lower on these than the other questions).

Question four asked students to imagine they are an amino acid in a folding protein and describe the intermolecular interactions they experience. Both the experimental and control activities contained embedded assessment questions of this type. Answers were scored out of two points; one point was given for a mention at least of the hydrophobic clumping effect, and the full two given only for a description that included interactions of hydrophobic amino acids with each other and with water, and interactions of water molecules with each other. Fractional scores were also given between one and two for partial explanations.

Students who had seen the dynamic experimental model scored higher on this question. Observing the model of molecules in motion may have made it easier to imagine the movement of an individual molecule, and its encounter with various charges.

Question five asked students to draw a picture of the folded shape of a protein in an oily solvent. The answers were scored out of two points; one point was given for a drawing that correctly showed the chain of amino acids bonded together and shaded according to their properties; two points were given if the chain was curled up into the correct shape: the hydrophobic region extended in a line and the hydrophilic region in a clump.

Predicting the behavior of a protein folding in oil is difficult; students for the most part reason about proteins in water, and the oily environment changes the rules. The use of a dynamic model that allows students to switch the solvent back and forth between water and oil may have resulted in a better understanding of this solvent effect. On a very similar question (question three), where the task is the same but the solvent is water, there was no significant difference between experimental and control groups. This may be due to the fact that the still pictures used in the control version were adequate to depict the folding process in water.

It appears that the activity with models better prepared students to answer those questions that required reasoning based on a mental model (as opposed to a set of facts). It's likely that the still images were more effective than models for supporting the learning style of some students, or for clearly conveying some of the concepts. The questions where the models did make a difference were the more difficult ones to do with visualizing and predicting the forces in protein folding. Constructing a mental model of protein folding, as opposed to memorizing a few factual statements about it, is a much more valuable form of learning because it enables students to reason about a broad range of phenomena.

Improvement in Molecular Reasoning

To determine further the role of experience with the models on student conceptual understanding, in the second year of testing a set of 75 randomly selected pre- and post-tests from upper- and lower-level classes were analyzed for inclusion of the vocabulary of inter-molecular interactions in their reasoning about biological phenomena. Because the Stepping Stone activities were done as supplementary materials to a full-year biology course, the impact of approximately ten activities on student learning (and, by default, their answers to the post-test questions) could most directly be measured by the inclusion of references to inter-molecular attractive forces.

Student answers clearly show that experience with the models influences their responses. In pre-tests, students included references to inter-particle interactions 15% of the time in their essay answers, whereas in the post-tests they included inter-particle interactions 57% of the time.

For example, in answering the first question of the assessment –"The instructions on the fertilizer (KCL, a salt like NaCl) for your plant requests that you add water with the fertilizer. Explain what happens to the fertilizer from the moment you add the water to the point when the nutrients enter a root cell."— a student could simply answer that the fertilizer would dissolve when the water is added and then, when placed in the soil, would move through the soil and be absorbed by the roots of the plant <u>OR</u> the student could include references to the relationship of attractive forces between polar water molecules and ions, describe the thermal motion of particles in the distribution of the fertilizer in the water, and describe diffusion and osmosis of water with ions into the root. The second type of answer relates more directly to the content of the MoLo activities and their influence on student reasoning.

(Answers to that specific question reveal that eight percent of the students analyzed had references to inter-particle interaction on the pre-test and 50% included references on their post-test.)

<u>Increasing number of core molecular concepts</u>

Student answers were analyzed for the appropriate inclusion of molecular concepts. A randomly-sampled classroom was scored blind for number of molecular-level concepts included in answers to several questions. At least 50% molecular-level concepts were included in the post-test than the pre-test.

In addition, in the formative pre-/post test:

- **Temperature and Molecular Movement:** 50% of students on the pre-test understood that, as temperature is raised, molecular movement increases (atoms and molecules move more rapidly). On the post-test, 100% of these students understood this concept.
- **Linking DNA to Proteins:** Students generally understood the relationship between DNA and mutations and the expression of the phenotype, but rarely included the molecular interactions (the resulting change in folding of the protein) in their responses. On the other hand, student ability to describe and draw the forces related to protein folding appeared to be fairly sophisticated.
- **Hydrophobicity:** Students were capable of understanding that hydrophobic molecules are attracted to each other more strongly than to water molecules and therefore tend to stay together. Students could observe such so-called water-fearing behavior of amino acids, and work with such a concept. On the pre-test, only 16% of the students could reason about how molecules with hydrophilic heads and hydrophobic tails might assemble in water. In the post-test, 75% of the students could draw such an assembly, and, of those, 45% invoked intermolecular interactions in the explanations.

Common Misconceptions in the Molecular World

Believing particles are only occasionally in motion is a common atomic-molecular misconception. In general, it appears that working with molecular models diminishes this idea. There remains, however, a residual concept:

Misconception: Particles are not always in motion. “The fertilizer for my plant is added with water causing random motion of particles. The particles from the fertilizer and the water move at different speeds. The substances come together as one mixture."

As students begin to integrate structure and force information, there were a few common mistakes, such as:

Misconception: Fewer bonds are weaker bonds. This may have been spill-over from the distinction often made in class between two and three covalent bonds, as on the project test we had drawn two and three van der Waals attractions and asked students which would come apart faster when heated. They gratuitously added the information that two bonds would separate more quickly because they themselves were weaker.

Misconception: Hydrophobic proteins actively repel water. We believe some portion of that problem is linguistic, and can be clarified; another portion of the problem is the difficulty understanding competing and collaborating forces.

Though the central dogma is taught in high school biology, our modeling activities on *Protein Folding* and *From DNA to Proteins* illustrate the forces and interactions among these molecules and the molecules that surround them. These activities, have also brought to light weakness in student understanding of the material.

Misconception: Nucleotides and amino acids are interchangeable. We believe this occasional difficulty sprang from the novelty of the protein material. Students needed more time and practice with proteins; we need to elicit their vocabulary more often.

Misconception: Amino acids can easily change from hydrophobic to hydrophilic and vice versa. This mistaken impression appears to spring from the modeling interaction of changing an amino acid’s characteristic without appearing to create a new polypeptide.

Misconception: Raising temperature “kills” proteins. While there were not many anthropomorphisms in student work, references to live enzymes were often elicited in response to a question about the effect of elevated heat on a system that included a working enzyme. “The enzyme would be killed.” While more informed students talked about denaturing and the subsequent loss of function, many other students used “killed” as metaphor, while others appeared to take the word literally.

Misconception: Whole genes are involved in all mutations. This is a fairly natural mistake, as we did not address the gene in our model, but simply a stretch of nucleotides.

but now need to work with teachers to clarify that mutations can involve as few as a single nucleotide pair.

Conclusions and Implications for Future Work

A. Conclusions

1. Our goal to create supplemental, stand-alone activities that can be inserted into biology classes was very successful. A sequential set of ten activities, some in two parts, were developed students use of molecular reasoning.
2. Materials were implemented in a wide variety of ways. The materials were used in a range of low-level to high-level classrooms. They were also used on a single computer as a demonstration, with individuals and pairs on computers, and as homework. The flexibility of the materials meant that they complemented most curricula and textbooks. For the most part, our Stepping Stone activities were on target (with the exception of two, weak attractions and proteins) for the high school curriculum and reading level. The exceptions were included because they covered critical content that we believed will soon be required of high school biology students.
3. We were particularly interested in the inclusion by students of molecular reasoning because it supports the ever-increasing molecular biological curriculum in high schools. Though the models were used infrequently in the classrooms, evidence of molecular reasoning on post-tests that reflected exposure to the dynamic models was present. This is impressive considering the minimal exposure to the models (5-10 class periods over the course of the entire year).
4. Whereas a few years ago, the content of the models would be considered outside the domain of the biology classrooms, teachers are seeing the merits of including more chemistry- and physics-based models that illustrate the dynamic interactions of atoms and molecules and how they might support biological phenomena. Such models are at an appropriate level for the students (as the models are not based on learning complex mathematical formulas). Of course, the inclusion of some of the models is still suspect for some teachers, as some content, e.g. protein folding, is still believed to be outside the domain of the biology classroom.
5. Teachers in the MoLo field test successfully implemented activities without much professional development other than the tutorial and a start-up workshop. They tended not to read the associated material provided by the database. Professional development could have a profound impact on the teaching and learning of this content and on developing a type of connections we refer to as "molecular logic." Teachers who could be trained to understand the connections among the activities, the logic that builds, and the connections to the other content would, we believe, have a strong foundation that could profoundly change the curriculum and in turn build student understanding.
6. Based on teacher interviews and email, it appears that a majority of the teachers who used the materials this year have recruited colleagues to use them next year.
7. We are developing a set of guidelines for creating activities using dynamic models that reflect what we have learned about design, pedagogy, and assessment with this

complex modeling world. As the software also incorporates an authoring system that is accessible by all users, we will draft recommendations for other users, too.

8. The capacity of *Molecular Workbench* to model DNA production of proteins, both translation and transcription steps, and mutations, was achieved.
9. The capacity of *Molecular Workbench* to embed assessments and pedagogy supports such as hints was achieved. Teachers could not use the assessments, however, unless they printed them out. Future work could be to develop a login system for teachers to access a database of their own students reports.
10. Deeper integration of these materials in physics and chemistry courses would make a profound background for a physics (first), chemistry and then biology curriculum.

B. Teacher Testimonials

Teacher testimonials from telephone interviews, email, and on-line feedback forms have provided insight into the implications of incorporating the model-based activities into the classroom. Here are a few quotes from the teachers that address different implications.

Teacher one speaks to the **retention** of the content over years.

> “I strongly agree that teaching with your materials has enhanced student understanding. For instance, this year in genetics we are just now discussing protein synthesis. Many of the students I have in genetics as juniors are the same ones I had last year in honors biology as sophomores - they are the ones who field tested the activities last year. I can tell they still remember the activities about protein self-assembly and it has made a difference in their understanding this year.”

Teacher two talks about how it helps provide **another resourc**e for going over content.

> “This material translates really well. I have noticed that my students appear to understand the content much better. I don’t have to spend as much time teaching and reviewing.”

Teacher three mentions that it **complements the curriculum**.

> “Last year I asked the project team to expand on the transcription and translation activity. They did it and I really love it. It also fell right in with the content. It was really easy to integrate.”

Teacher Four describes how **additional curriculum is critical** for understanding biology these days.

> “I honestly have never linked van der Waals attractions in the past. It is not traditionally in our curriculum. It was a challenge for me. On the other hand, more emphasis on hydrophobicity and polarity with biological applications was critical and really helped a lot.”

Teacher five describes how **visualizations help students understand** otherwise complex materials.

> “This year has been very good for my students to use the activities. To see the models. One girl kept going ‘Oh wow I see how it works, So like the air molecules are hitting our skin right now?’ That was something that wouldn’t have happened. The models went a long way towards helping them to apply the content to real situations.”

Teacher six adds to this:

> I thought that these simulations deepened their understanding; and for those who are kinesthetic/visual learners, it made the difference between memorizing and knowing.

Teacher seven describes **challenges of incorporating materials**.

> “I like the fact the kids can see what is going on. For example the activity on the biomolecules really helps them to see what things look like. It is much more efficient use of time than the traditional ball and stick activity. Or, for the weak attraction activity, though it was very long, towards the end it tied things to biological examples. I don’t normally teach weak attractions to the degree that activity took it to. That was a bit of challenge for me to keep it connected to biology and consistent with the biology content requirements. I can see why it is important but there was a tension.”

VI. Project Assessment and Oversight

Dr. Sigmund Abeles, an outside evaluator, performed project evaluations. He visited a sample of classrooms, reviewed project materials, met with project staff, and submitted his findings to both the staff and the oversight officers of NSF.

The Molecular Logic Board met twice, one face to face and one via on-line meeting and phone. Their counsel was heeded, with materials becoming more focused and select. Select members of the board served as scientific advisors, particularly Dr. Ringe from Brandeis University.