RESEARCH REPORT

Science ASSISTments: Tutoring Middle School Students' Inquiry Skills in the Domain of State Change

An Interactive Qualifying Project Report

Submitted to the Faculty

Of the

Worcester Polytechnic Institute

In partial fulfillment of the requirements for the

Degree of Actuarial Mathematics

By:

Steven DiTullio

October 25th, 2010

Approved By:

Professor Janice Gobert

Abstract

This paper outlines the development effort and a small pilot study conducted to assess students' understanding of state change in order to: 1) characterize middle school students' inquiry skills, 2) develop an effective set of tutoring prompts for inquiry within a technical environment, and 3) determine the effectiveness of various prompts and scaffolding tasks at improving inquiry skills.

Table of Contents

The ASSISTments program[1] was developed by Computer Science Assistant Professor Neil Hefferenan of Worcester Polytechnic Institute (WPI) and his colleagues at Carnegie Mellon University (CMU). The system was created as an "intelligent tutoring" web-based system that provides example problems and reacts to inputs from the students in real time. Originally developed to provide math tutoring for kids, Math ASSISTments is presently used by 3,000 students and 40 teachers in Massachusetts. The goal is to improve each student's math learning. One of the measures to determine the success of the ASSISTments system is the Massachusetts Comprehensive Assessment System (MCAS). There are several mathematical concepts covered in the ASSISTments problems that are incorporated to improve a student's MCAS scores.

The ASSISTments system is designed with three separate methods of guiding students through each mathematical problem; hints, buggy messages, and scaffolding. Hints and scaffolds are only provided at the request of the student when they are struggling to answer a particular question. Each hint or scaffold provides more and more information towards the correct answer. The difference between hints and scaffolding within the ASSISTments system is that hints are directed at the specific problem the student is working on while a scaffold allows the student to see outside the problem at hand by breaking down the concepts of a complicated problem in order to simplify the question. The buggy messages which are the third method to guide students, only pop up when the student provides an incorrect answer. Since the student's lose some credit when they utilize the hint option, they do not overuse it.

Students and teachers both interact well with the program. The math ASSISTments system has become part of the weekly routine at many schools and the students learn math while utilizing the learning environment. The students are also encouraged to provide comments for the software developers so that the program can be tailored to their particular needs. The system also allows teachers to track each student's grades and see their progression as they use the system.

Development of the Science ASSISTments program

The ASSISTments program was extended to include a system to help students with science, specifically Physical Science, in 2007. Janice Gobert, a professor at WPI, lead a team of graduate and undergraduate students to develop the new system. Professor Gobert's expertise is Cognitive Psychology as well as Science Education. She co-directed the Modeling Across the Curriculum project[2] where she and her team developed individual micro-worlds which allowed students to explore scientific phenomena in the areas of Biology, Physics, and Chemistry. Her expertise combined with the technical background in Computer Science of professor Hefferenan and his ASSISTments program form a team dedicated to improving inquiry skills in middle school students and promote a better learning environment in the classroom.

In its current form the science MCAS test assesses only the content knowledge of the student and it does not measure inquiry performance. This project aims to provide measures to determine a student's inquiry skills and as such, this effort represents a significant advancement in science assessment. There have been two approaches to date that have been taken to assess students' inquiry skills. Short answer tests of specific inquiry skills (cf., Alonzo & Aschbacher, 2004) have been used in the past as a method to learn about a student's inquiry skills. Short answer tests can be incorporated into large-scale assessments. It is unclear however whether these tests properly identify inquiry (Black 1999; Pellegrino 2001). The other test type used is a paper and pencil performance assessment. These performance assessments are more authentic because they require the student to have specific skills to solve real problems (Baxter and Shavelson 1994; Ruiz-Primo and Shavelson 1996). This fact along with the high costs associated with conducting an inquiry assessment test is why inquiry is seldom assessed in today's schools.

The Science ASSISTments project team will respond to state needs, meaning their aim will be to develop inquiry measures for science as well as scaffolds for students geared to teach them inquiry skills. The overarching hypothesis is that teaching the students inquiry skills will help them gain a deeper understanding of science content. The Science ASSISTments team will use prior MCAS exams as well as inquiry items developed in house to address the learning of inquiry skills. This paper is focused on the development of background for the development of inquiry scaffolds for the science ASSISTments program in the topic of State Change.

Why Teach Inquiry?

Inquiry is an important skill when a student attempts to perform logical scientific experiments as well as make decisions in real life scientific problems. An example of this would be testing the effects of different variables on the outcome of a particular experiment. The skill, called control for variables strategy, does not come naturally to most individuals and thus many students are lacking these important skills (Kuhn, 2005).

Also, due to the large amount of time and effort it takes to teach and assess inquiry these skills are not often taught in a typical class curriculum. The definition of inquiry skills are as follows (NSES 1996):

Hypothesizing: making predictions about the outcome of an experiment

Conducting an experiment: developing a method of testing the hypothesis

Collecting data: correctly recording findings while performing the experiment and setting up tables of variables

Mathematizing: performing calculations and writing equations to formally describe collected data

Interpreting data: forming explanations that are consistent with data collected

Communicating: articulating the steps of inquiry (describing results, etc.)

The long term goals of the ASSISTments project is to improve the inquiry skills of middle school students so that they can transfer this knowledge to other topics in science, and ultimately to real world problems and improve upon content learning as well. Research shows that students are greatly benefited by learning these skills at the middle school level (Schunn, Raghavan, & Cho, 2007). How feasible it is for students to take the inquiry skills they have been taught and apply them to new domains will be determined by the outcomes of tests and experiments that are to be performed during the Science ASSISTments project.

Adapting the Math ASSISTments System for New Science ASSISTments project

(Gobert, J., Heffernan, N., Ruiz, C., & Kim, R., 2008)[3]

The Math ASSISTments system is a server-based (thin client) tutoring system implemented using Ruby on Rails, JavaScript, and XHTML/CSS style sheets. There are several advantages to this server-based approach. Using this model the software will be run entirely through a web browser and thus eliminating a need for the user to install anything onto his or her machine. The time and cost associated with installing software onto machines in a computer lab is virtually eliminated. This also provides us with a greater control over the distribution of the product since we can control what new content becomes available as well as when it becomes available for all users regardless of their location. Updates and configuration to the program become much more manageable using a web based system. This approach is also gathers the data collection on a per student basis in a central location as opposed to on each individual student's machine. Finally, running in a browser abstracts from creating distributions for different operating systems, thus supporting Windows or Mac users. There are some disadvantages to running the program on the web, responsiveness being the most prevalent.

The ASSISTments architecture is able to combat these disadvantages in multiple ways. The system is housed using Mongrel servers, which is a load balancing cluster with several computers and multiple application servers that distributes the processing workload. This allows us to theoretically scale up the software to support more users by adding more computers with application servers. Additionally, smart indexing on frequently used values in the database drastically reduces the time of responding to user actions. Most importantly, the implementation pushes more of the processing onto the client.

The servers are responsible for on demand content distribution and log gathering but push data processing tasks onto the client. For example determining the correctness of an answer is handled by the client. This greatly reduces the load on the database and servers since the client does not need to constantly query back to the server for as much. At the start of an assignment, the client requests and stores questions asynchronously while the student begins working. Network or server delays are hidden by the asynchronous communication and subsequent questions are streaming in the background. Since the content is stored on the client, the progression through an assignment is handled by the client with only log messages sent back to the server.

Running the Science ASSISTments program client side creates many constraints that we must overcome in order to create an interactive environment. These environments must download quickly, be responsive, be reusable so we can create multiple kinds of questions involving each one, must be portable, and most importantly these environments must be able to log user actions as well as answer questions. Most browsers do not natively support the idea of a "canvas" whereby on can author an animated, graphical environment. This makes authoring such environments very difficult. For success, we require the machines be equipped with a third party environment such as Flash.

The overall design goal is to equip the ASSISTments system to handle questions involving interactive environments. This enables the project team to reuse the existing framework and infrastructure that manages both the student and teacher accounts and lets the team create and assign problem sets on a per environment basis. This also affords us to send streams of problems to the client behind the scene and asynchronously log student actions and their results. There are two ways in which we will extend the system. First the team needed to extend the problem framework to allow Flash objects to appear inside the problems of each environment. In previous iterations only static images could be displayed inside the problems. Second the team enhanced the tutor component of the program to interpret the results from the Flash objects and log any actions the student takes. Another benefit is that the existing rendering system with all of the correct/incorrect graphics is reused as well.

Computer Micro-worlds

Science ASSISTments differs from Math ASSISTments in that most scientific problems tend to be open-ended. In Science ASSISTments, students are given a specific goal and time to explore with one of the micro-worlds. This creates the need to develop an interactive micro-world that the student can utilize and manipulate in order to solve each problem. These micro-worlds are designed in accordance with previous findings from Professor Gobert's previous work on the MAC project (mac.concod.org). What happens is the student is given a specific task, for example: "To maximize the speed of the orange ball after the blue hits the stationary orange ball at a speed of 4 m/s". The student then runs the experiment by changing the mass of the two balls to see the effect that mass has on the final velocity upon collision of the blue ball with the stationary orange ball.

The Current Curriculum Model

The structure of a curriculum plays a big part in student development. Creating a well balanced curriculum that establishes a powerful learning environment for students is always the goal. This is achieved by carefully outlining not only what subject matters teachers will cover, but how to best present these ideas to their students in order to minimize the prevalence of misconceptions. It is equally important to plan out how you are going to teach a particular content area.

There exist guidelines formed by agencies in the US to lay down a foundation for what a typical curriculum should have. These are broken down by subject and grade level and are used as a guideline for instructors to build around. Some of these agencies even go so far as to suggest different activities that could be done in a classroom setting in order to enhance the child's understanding of a topic. For the purposes of discussion, this paper is going to focus on the middle school demographic and the subject of science.

One agency that involved in the creation of curriculum guidelines is the National Science Resource Center. This agency was created by the Smithsonian Institution and the National Academies. The Smithsonian Institution was founded in 1835 as a gift to the United States of America from James Smithsonian. Its goal is to forward the "increase and diffusion of knowledge…"[5] The National Academies perform a unique and very similar public service by bringing together committees of experts in all areas of science and technology. These experts serve pro bono to address critical national issues and give advice to the federal government and the public. These two institutions provide a strong platform for the NSRC to branch off from. The main purpose of the NSRC is to improve the learning of students from grades K-12 not only in the US but throughout the world.[4]

The NSRC is an intermediary organization that attempts to bridge the gap between researches on how students learn with the best practices in the classroom. From this research the NSRC creates a series of guidelines to follow when building a curriculum for a particular grade level. It then becomes the responsibility of the teachers and faculty of a given school to build a curriculum in accordance with these standards and guidelines. The NSRC has become a leader in the field of science education development by:

- Building awareness for science education among leaders;
- Helping develop science education leadership among groups from school districts and states;
- Conducting programs to support the professional growth of teachers; and
- Developing and disseminating information about exemplary science instructional materials.[6]

Following this model, the NSRC has developed standards for both instructors and students as to what material should be covered in each grade level as well as some guidelines for how to best execute each lesson.

The NSRC Standards

The following list can be found on the NSRC website. It summarizes the content and process standards addressed in *Properties of Matter* subject. These students are typically in grades 5-8 range.

Physical Sciences (Properties and Changes of Properties of Matter)

- Any substance has characteristic properties that can be tested and measured and are independent from the amount of substance that is collected for measurement. These properties include density, a boiling point, and solubility. When dealing with a mixture of several substances it is often the case that the individual substances can be isolated and removed from the mixture using these properties.
- When a substance is involved in a chemical reaction with one or more other substances, it reacts in characteristic ways. This process changes the substances involved to form a new substance with different characteristic properties. Mass is not lost during this process. Different substances are typically placed in different categories in order to separate them by their characteristic properties. For example, metals are grouped together because they all have similar properties.
- Under normal laboratory settings, a chemical element will not break down in experiments involving heating, exposure to electric current, or reactions involving acids. There are more than a hundred known elements that combine in a multitude of ways to produce compounds when those elements are involved in reactions. These elements and the compounds they form are what make up all the living and nonliving substances we know of.

While these guidelines a good starting point for discussion about what a curriculum should include, science does not exist in a vacuum. In order for the students to understand the inner workings of chemistry they must be exposed to the proper tools for measurement. These tools vary in age, application, and practicality however they are all common technologies used in real world problem solving. This opens the door for a discussion on how these different technologies came to be, what they are today used for, and how technologically advanced they are.

Because technology and science are so closely aligned with one another, it is imperative that the root and purpose of each of the different technologies used in class be properly explained to the students. Below is a list of guidelines the NSRC dealing with teaching students the relationship between technology and science.

Science and Technology (Understandings About Science and Technology)

- Science and technology go hand in hand. Scientific advancement brings about a need for more sophisticated instruments and leads to the development of new technologies, or advancement of older ones, through technology. Technology makes scientific discovery, analysis, and progress to continue. By providing the instruments and techniques that better enable observation of phenomenon that would otherwise be unobservable technology provides us with the tools for investigation, inquiry, and analysis.
- While in the classroom examples are "clean" and work out exactly, real world problems do not have perfectly designed solutions. All technological solutions involve trade-offs with regards to safety, cost, efficiency, and appearance.
- All technologies have limitations with regards to accuracy. Some of these constraints are totally unavoidable, for example the properties of certain materials or the effects that friction or weather might have on an experiment. Other limitations limit the choices in design of the technology itself with regards to

durability in the environment the technology needs to work in, the aesthetics of the device, and the safety of the people operating it.

These are the topics that the NSRC has decided are important to be included in a chemistry curriculum during anywhere from grade 5 to grade 8. However before teachers can use this information to build a curriculum for their classes it is important to understand how the children think about different topics in chemistry. This is done by analyzing the student's inquiry skills. Using data collected on the inquiry skills of students in grades 5-8 it becomes possible to not only provide guidelines to what materials should be covered but also on the best method of teaching as well. Below is a list of the guidelines created by the NSRC.

Skills necessary for students to conduct scientific inquiry

- Students should be able to recognize questions that are best answered through scientific investigations. They should be able to take questions that are broad and ill-defined and refocus them more clearly. One very important aspect of this process is the ability to clarify questions and be able to relate them to objects and phenomena that can be described, explained, or predicted in a scientific investigation. Students should be able to identify questions that specifically deal with scientific ideas, concepts, and quantitative relationships that guide the scientific investigation itself.
- Students should be able to design and carry out a scientific investigation on their own. They should be able to develop general skills such as systemic observation, identifying and controlling variables, and making accurate measurements. They

also should develop the skills to clarify the ideas that are effecting and guiding the inquiry and to understand how these ideas are related to current scientific knowledge. Students should be able to formulate questions, design investigations, execute them, and interpret the data they collect during the investigations. The students also should be able to find evidence in the data to support their explanations, propose alternative explanations, and critique these explanations they have found and the procedures they have followed.

- Students should be able to identify and utilize the proper tools and techniques to gather, analyze, and interpret data. The use of these tools and techniques, including the quantitative skills to calculate and analyze data, will be guided by the original question asked and how the student designs his or her own investigation. The use of technologies such as computers for collection, summary, and display of evidence is a part of this standard. Students should know how to use these machines to access, gather, store, retrieve, and organize data by using software and hardware designed for these purposes.
- Students should be able to describe, explain, predict, and model their findings by using evidence discovered during the investigation. They should base these findings on what they observe. As they develop the cognitive skills required they should be able to distinguish between their explanations and the description of what's going on. The students should also be able to identify relationships they uncover on the basis of evidence and logical argument. This standard requires the development of a subject matter knowledge base so the students can effectively conduct their own investigations. By developing their own explanations, students

can make connections between the content of science and the contexts within which students develop new knowledge. In other words, this helps remediate misconceptions students may form based on false assumptions.

- Students should have the skills to think critically and logically in order to identify the relationships between evidence and explanations. Thinking critically about evidence includes the students deciding what evidence should be used and accounting for data that is irregular in a given situation. More specifically, students should be able to look over the data that is collected in a given experiment, summarize it, and formulate a logical argument about the cause-and-effect relationships that exist in the experiment. They should also be starting to state some of these explanations in terms of the relationship between two or more variables that exist in the experiment.
- Students should be able to recognize alternative explanations for a given phenomena and analyze these predictions. They should be able to both listen to and respect the explanations proposed by their fellow students. They should be open to the idea that people will have different explanations and conversely they should be able to accept criticism and skepticism of other students. They should also be able to consider alternative explanations.
- Students not only need to be able to develop scientific procedures and explanations, but they should also be able to communicate them effectively to others. In time the students should be able to communicate experimental methods, follow instructions provided for a given experiment, describe the

observations they find, summarize the results of other groups, and tell other students about their own investigations and explanations in a competent fashion.

- Students should be able to recognize the role of mathematics in all aspects of scientific inquiry. Mathematics is essential to asking and answering questions in and about the natural world and it is imperative that students understand this concept. Mathematics can be used to ask questions, gather data, organize results, and construct convincing explanations and present evidence in a logical manner.

Understandings About Scientific Inquiry

- The questions asked about specific phenomena are what drive different kinds of scientific investigation. Some investigations involve the observation of objects, organisms, or events and describing each of them in turn. Others involve the collection of specimens for research or experimentation. Some studies will require the student to seek out more information or the discovery of new objects and phenomena while others are more hands on and involve the construction of models.
- The current scientific knowledge available and the understanding of that knowledge guide scientific investigations. Different scientific domains use different methods of investigation, different core theories about the domain, and different standards to advance scientific knowledge and understanding.
- Regardless of what the inquiry is about, mathematics is important in all aspects of scientific inquiry.

- Technology used to gather data enhances accuracy and allows scientists to analyze and quantify results of investigations. The more advanced these technologies become, the more accurate the data becomes and the more concrete evidence becomes.
- Some characteristics of scientific explanations are that they emphasize evidence, they have logical arguments, and they use scientific principals, models, and theories. The scientific community will continue to accept and use these explanations until they are displaced by better ones. This process is known as scientific advancement.
- The only way science advances is through legitimate skepticism. Asking questions and challenging other scientists' explanations is a big part of scientific inquiry. Scientists go about doing this is by evaluating the explanations other scientists provide during their experiments, identifying evidence where in the logic can be proven faulty, pointing out new statements that go beyond the current evidence, and suggesting alternative explanations for the same list of observations.

The NSRC has laid out guidelines for building a chemistry curriculum for students in grades 5-8 using three main building blocks. The first notion is to build up the knowledge of the subject itself.

Science and Technology Concepts for Middle Schools

Science and Technology Concepts for Middle Schools (STC/MS) is an inquiry-based middle school science curriculum developed by the National Science Resources Center.

This student and teacher site supplements the STC/MS curriculum and can be used as a research guide for students.[7]

STC/MS is an 8-module, inquiry-centered, middle school science curriculum developed by the National Science Resources Center. Each STC/MS module provides opportunities for students to experience scientific phenomena firsthand. A comprehensive, research-based curriculum, STC/MS is aligned with the NSES of the National Research Council (NRC).

Each STC/MS module is based on a four-stage learning cycle that is grounded in educational research and practice.

- First, students focus on what they already know about a topic.
- Next, students explore a scientific phenomenon or concept, following a well-structured sequence of classroom investigations.
- Third, students reflect on their observations, record them in science journals, draw conclusions, and share their findings with others.
- Finally, students apply their learning to real-life situations and to other areas of the curriculum.

The STC/MS program builds on the skills and knowledge developed in the STC® curriculum, with content balanced among the life sciences, earth sciences, physical sciences, and technology. Each module is developed by teachers, scientists, and evaluators and is field-tested in urban, suburban, and rural classrooms nationwide. The materials are professionally evaluated and reviewed by an advisory panel of teachers, scientists, and science educators before publication. [5]

Alignment of STC® and STC/MS™ Science Curriculum Modules

Grade Level		**Life and Earth Sciences**		**Physical Science and Technology**	
STC	K–1	*Organisms*	*Weather*	*Solids and Liquids*	*Comparing and Measuring*
	2	*The Life Cycle of Butterflies*	*Soils*	*Changes*	*Balancing and Weighing*
	3	*Plant Growth and Development*	*Rocks and Minerals*	*Chemical Tests*	*Sound*
	4	*Animal Studies*	*Land and Water*	*Electric Circuits*	*Motion and Design*
	5	*Microworlds*	*Ecosystems*	*Food Chemistry*	*Floating and Sinking*
	6	*Experiments with Plants*	*Measuring Time*	*Magnets and Motors*	*The Technology of Paper*
STC/MS	6–8	*Human Body Systems*	*Catastrophic Events*	*Properties of Matter*	*Energy, Machines, and Motion*
	6–8	*Organisms—From Macro to Micro*	*Earth in Space*	*Light*	*Electrical Energy and Circuit Design*

Note: All STC units can be used at one grade level above or below the level indicated on the chart. STC/MS units can also be used at grade 9.

[7]

The modules can be sequenced for 2 one-year courses, each year consisting of a module from each of the 4 science/technology strands, or as 4 one-semester courses for earth science, life science, physical science, and technology.

Research and Development

The STC/MS modules keep inquiry at the center of the learning process. Each module was created through a rigorous research and development process. Scientists and educators, including experienced middle school teachers, acted as consultants to teacher-developers, who researched, developed, and trial taught the units, then coordinated field testing and produced the final STC/MS units.

STC/MS was evaluated by the Program Evaluation and Research Group of Lesley College, located in Cambridge, Massachusetts and the Center for the Study of Testing, Evaluation, and Educational Policy, located in Boston, Massachusetts. The final editions of the units incorporate teacher and student field-test feedback and technical reviews by leading scientists and science educators who serve on the STC/MS Advisory Panel.[7]

7

Developed for Grades	6–8	6–8	6–8	6–8	6–8	6–8	6–8	6–8
Module Titles	Human Body Systems	Catastrophic Events	Properties of Matter	Energy, Machines, and Motion	Organisms From Macro to Micro	Earth in Space	Light	Electrical Energy and Circuit Design
Science as Inquiry								
Abilities necessary to do scientific inquiry	●	●	●	●	●	●	●	●
Understandings about scientific inquiry	●	●	●	●	●	●	●	●
Physical Science								
Properties and changes of properties in matter		●	●					
Motions and forces				●		●		
Transfer of energy		●		●		●	●	●

STC/MS™ is an inquiry-based middle school science curriculum developed by the National Science Resources Center. This student/teacher site supplements the STC/MS™ curriculum and can be used as a research guide for students.[5]

Student Misconceptions

As with any subject matter in school, not all students clearly understand the larger concepts in a subject matter. Regardless of grade level or skill students form misconceptions in order to fill in the blanks and complete their understanding of the topic. These misconceptions are caused by several different factors.

The structure of the curriculum is one of the reasons for students to develop misconceptions. Often subjects are only discussed and not addressed with the students in order to cover all the material the teacher is required to complete. This leads to students questioning what is really going on in a particular subject. For example, when students are taught about phase change in class, their instructors talk about different terms involved with the subject such as evaporation. They give a definition of the term and then move on to talk more in detail about the larger subject matter of phase change. Students will inevitably ask any questions that come to mind in an effort to better understand the inner workings of evaporation. In the interest of time an instructor cannot answer every question that the students bring up and this leads to the students coming up with their own explanations in order to make sense of what is going on.

Some subjects are inherently far more difficult for students to grasp than others. For example, students as a whole have trouble with certain ideas about the chemical equilibrium. When dealing with combustion, students have little trouble understanding what is going on. They not only can visualize what is happening in a lab setting, but they also have experience with combustion in their own lives. This exposure leads to an easier time conceptually for the students. However when the concept of the chemical equilibrium is brought up and they are told that matter is conserved in the system, students start to have problems. They look at the reaction and see that the amount of fuel they started with appears to be smaller than the amount of residue that remains in the aftermath. Students have even more difficulty when talking about subjects they do not have personal experience with.

The students' own construction of knowledge is another source of difficulty when it comes to forming an in depth understanding of a subject matter. Students try to make connections between what they are taught and what they already know to be true. This by itself is not an issue and should be encouraged by instructors, as it will in fact help students to better grasp concepts. However the problem occurs when what the students know to be true is in fact a misconception. When told about heat energy, students are very comfortable with the notion that when some object is subject to heating, the object's temperature will increase. However where the students then have difficulty is in addressing "what happens when the heat is taken away?" Students are aware that the object will cool down but are unsure as to why. As a result, they use a logical approach and say that if there is a force that is heating the element there must exist some force that in tern cools it back down when the heat is gone. When these connections are formed on misconceptions, often it makes correcting the misconception far more difficult because the students have now built upon that foundation with other ideas that are also incorrect because they are also based upon misconceptions.

There have been countless studies done on the misconceptions students form during learning science. The concept of phase change is often one of the most dealt with as it is intrepid that students understand this concept in order to better explain later topics in science such as thermodynamics, heat transfer, evaporation, and several other concepts dealing with states of matter. Yet despite the importance of the topic, students typically have a large degree of difficulty truly comprehending the ideas involved in phase change.

The article "Children's Conceptions of Gas" was written by Ruth Stavy and ran in the International Journal of Science Education in 1988, volume 10, issue 5. The article is about a study conducted in Israel dealing with the topic of phase change. More specifically the study was an effort to discover children's misconceptions about gases in particular.

The study was conducted with a group of students in the 4th through the 9th grade level. The goal was to isolate why students had certain misconceptions about the gas state of matter and where the misconceptions stem from.[8] The reason for choosing such a wide age range was so they could test the curriculum itself as well. Students in the 7th grade level are first taught about the states of matter. Therefore the study also shows how well the students understand the states of matter in the 7th grade and beyond as well as identify how well the students can understand the idea of phase change without formal teaching.

How they went about this was they asked the students questions about what was happening during the creation of soda water. The pupils were shown a CO_2 container before the container was emptied and after it was emptied. The students were then asked questions about its weight. Students were also asked about the water itself and how, if at all, the weight had changed now that it was the carbonation. The student's responses were then documented and categorized by how well they understood the mechanisms of gases and what was going on in the given scenarios.

Student in general had a much easier time describing and answering questions about the soda water. Most students responded that the water must weigh more with the gas now mixed in because, “the gas has a weight that’s been added to the water, and therefore it must weigh more than before.”[9] The few exceptions typically said there was no change because gas has no weight. The other incorrect answer was that the water is now lighter than before because the gas bubbles have now added their lightness to the water because gas is lighter than liquid.

When asked about the CO_2 container students were asked whether or not the tank weighed as much as it did, more than it did, or less than it did now that the gas had been released. Students had more trouble with this question as shown in the data. The most typical response was that the container’s weight had not changed because “gases have no weight”. This trend was found in students until they reached the 7th grade level where the more common response became, “the gas was in the container, now it’s gone, so the container must weigh less because the air has a weight”.[9] The difficulty with the CO_2 container scenario is thought to be because the students cannot see the gas escaping the container as they can with the bubbles in the water. Without the visual aid the students have a much more difficult time understanding the properties of CO_2 gas.

As a final task, all of the students were asked to provide a definition for a gas. Up until the grade 7 level the majority of students would define a gas by using examples they were familiar with such as cooking gas. They did not show much understanding of the properties of gases. Some of the younger students and the vast majority of the 8th grade kids provided a definition that described gas as a state of matter, without any reference to the movement of molecules. Only when you got to the 9th grade level did students start to use the particle theory of matter to describe gases. The following chart details the results below as a percentage of the students that provided a given response.

Table 1. Children's definitions of the term *gas* (per cent).*

Definition	Grade 4	5	6	7	8	9
By example such as:	85	80	65	35	—	—
Cooking gas	40	20	35	—	—	—
Gas in light drinks	35	15	25	—	—	—
Air	5	40	20	35	—	—
Steam	—	10	—	—	—	—
Others (tear gas, poisonous gas, laughing gas, etc.)	10	5	—	—	—	—
As a form of matter relating to:	—	20	20	45	75	35
Properties	—	10	10	25	5	—
State of matter	—	10	10	20	70	35
By means of the particulate theory of matter relating to:	—	—	—	—	(1) 25	(2) 80
Distance between particles	—	—	—	—	25	70
Motion of particles	—	—	—	—	10	20
Arrangement of particles	—	—	—	—	5	5
Attraction forces between particles	—	—	—	—	5	5
Not known or irrelevant	15	—	—	—	—	—

9

The study shows that students do not inherently come up with the idea of phases change without schooling. Once taught in the 7th grade about the states of matter and particulate theory the students begin to understand the nature of gases. Within the next year the students start to describe gases as a state of matter culminating in the 9th grade where students finally begin to describe gases as particles that are moving further and further apart. Student's ideas about phase change are also not consistent when the scenario is changed. Their reactions to essentially identical problems are different and only when the students reach grade 7 do some of these misconceptions begin to be remediated. For example, some students think that the gas has no weight in the water but does in fact have weight when it is in the cartridge. Finally the student's notions of the particulate theory of matter are very fragmented and are not applied in every situation properly.

With regards to phase change, the gaseous state is often one that provides the most difficulty. Even in the simple case of using water, students have a hard time identifying the properties involved with water vapor or steam. This is illustrated from their lack of understanding with regards to the process of evaporation.

In the International Journal of Science Education there was a special issue in 1989, volume 11, with a study on children's ideas about evaporation. The study was conducted by Terry Russell, Waynne Harlen, and Dot Watt collaboratively with Liverpool University in the United Kingdom. The products of the study were to gain a great understanding of children's ideas about evaporation, develop some ideas about student's conceptions, to develop techniques that were both effective and economically feasible for a classroom setting, and the identification and creation of several intervention techniques to aid in future teaching.[10]

During the study the students were selected from 6 schools in England taking a sample of 400 students varying in ages from age 5-11. These students were interviewed individually for 35-40 minutes both before and after they had completed the classroom activities created to help them understand evaporation. The classroom activities were broken down into three sections.

First there was an exploration phase where students were provided various thinking activities in the classroom setting. These activities were used in parallel with the student's normal classroom activities. Participation was not made mandatory for the students however they were made aware of what was going on. The children's ideas were then analyzed more carefully in the intervention phase of the study. During this phase the study moved from casual observation and collection of the student's ideas into an attempt to optimize the child's thinking and reasoning about the activity. There were three different activities the students were asked about: monitoring water level in a transparent 12L tank, monitoring wet clothes left to dry and monitoring sugar in coffee solutions in saucers. There were six basic strategies used to guide the students during each activity:

1) Encourage children to test their ideas on the assumption that a scientific test might lead children to consider some ideas unsatisfying.
2) Consider vocabulary carefully; vocabulary work was always urged to be with reference to real objects and events.
3) Encouraging generalizations from one specific context to another, which children themselves identified.
4) Finding ways to make imperceptible changes perceptible.
5) Testing the "right" idea along with the child's own ideas.
6) Articulation or criticism of ideas as a group or a class.[10]

The final step was the post-intervention elicitation. Pre and post interview data were summarized. This allowed the data to be broken down by age group and showed which ideas were consistent in each age group. Each activity was also broken down according to the different responses.

The evaporation of water from the transparent tank elicited 3 different responses from the students. Some students argued that the water was now simply gone. There was no explanation given as to where it now is or how it got there. This group of students showed a lack of familiarity with the conservation of water idea that the water cannot just be destroyed and it must now go somewhere. Typical responses were "the water is now gone", "it disappeared", "it's dried up". This group was almost entirely populated by the younger students.

The next group of students made note that the water has changed location but has not changed its physical form. Students attributed this movement to one of three devices. The water was either moved by some known or unknown human or animal agent, there was a non-animal physical agent that was moving the water which was typically attributed to the sun, or the students said the water had been moved by the agent into another medium such as moved by the sun into the clouds. The first view was typically popular with the younger demographic and the other two grew in popularity as the age of the students being tested increased.

The final classification of the students was the group that attributed the movement of the water to a change in the water's physical state. Some students argued that the water had been relocated in some perceivable form (water droplets, mist, steam) while there were others who said that the water was not seen (water vapor, gas). The students drew pictures of what they meant.

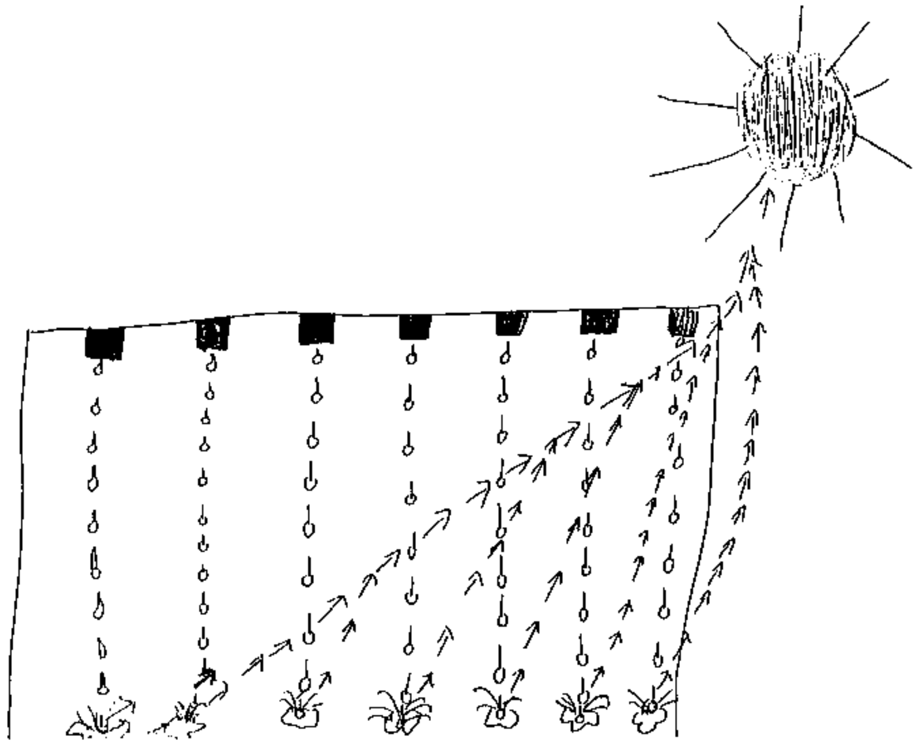


Figure 1. Child's drawing of relocation of water from clothes hanging on a line to the site of the agent.

10

Figure 2. Drawing of relocation of water to a site other than that of the agent (no physical change in the nature of the water).

10

This argument about water traveling from the container to a different medium by changing its physical state was popular only with the older students and is as far as the students can be expected to get on their own with the concept of evaporation. Below is a chart detailing the breakdown of the student body tested:

	Infants *n=17*	*Lower juniors* *n=18*	*Upper juniors* *n=23*
No response	12 (2)*	—	—
No necessary conservation: focus on remainder	35 (6)	—	13 (3)
Human or animal agent	18 (3)	—	—
Change of location to site of agent	29 (5)	22 (4)	4 (1)
Change of location by physical agent	6 (1)	33 (6)	57 (13)
Physical change in water, but perceptible form	—	28 (5)	17 (4)
Water changes to imperceptible form	—	17 (3)	9 (2)

[10]

The other two activities where the water was drying from a wet shirt and the evaporation of both a coffee and sugar mixture were not as clear cut for the students. The water drying from the shirt added the possibility to the students that the shirt had absorbed the water in order to dry it. The process of evaporation seemed to be confounded by the functionality of a paper towel or a mop in the student's mind. The coffee solution as well as the sugar solution elicited the response from the students that the water in each case had evaporated and the solids had not. The other result mirrored that of the first activity.

Evaporation has been a difficult part of phase change for many students to grasp. In the International Journal of Science Education in 1994, volume 16, no 2, another study was conducted by Bar and Galili in Israel to further delve into the topic. Their goals in the study were to answer some common questions about how young children understand the process and properties of evaporation. These questions were:

1) Could children's views about evaporation be scaled according to age?
2) What is the relationship between these views and the children's understanding of both the conservation of liquid and the structure of air?
3) What is the effect of the study's testing method on the children's views?
4) Are children's views on evaporation context dependent?

The study was broken down into four phases of questioning. The first phase students were only asked, "What happens to water when it is spilled on the floor". In the second phase students were given activities to do dealing with conservation, evaporation, and air composition. For the conservation activity the students were presented with two equal containers of water. The contents of one container were then poured into a taller and narrower one and the students were asked if the amount of water was still the same. For evaporation the students were asked not only what happens to water that is spilled on the floor, but also where is could now be found. A similar question would then be asked by replacing the word floor with ground. The air activity was taking a piece of paper and waving it at the child. The student was then asked what causes the sensation of wind. It should be noted that phases one and two were done with students that have no yet had experience in classes with the concept of phase change. This was to asses if the students could grasp phase change without formal training in order to determine if the student's misconceptions come from outside the classroom. Also, the questions on air were only asked of students in at least the 4th or 5th grade as typically understanding of air composition does not begin until that level.[11]

Phase three consisted of a written multiple choice test, based around the same logic as the first two phases. The main difference was now the older students that had been exposed to phase change topics in school were now included in the test. Phase four focused on a very small group of students who during the phase 3 test had chosen the response, "water disintegrates into hydrogen and oxygen" to explain the evaporation process. These students were asked to explain how they came up with this notion. The table below breaks down each phase by age group.

Table 1. The research sample: age groups and numbers of subjects in each group.

Grade	*Age*	*Phase I*		*Phase II*	*Phase III*		*Phase IV*
Kindergarten	5·3	–	–	20	–	–	–
1	6·5	30	23	20	–	–	–
1	6·75	30	–	15	–	–	–
2	7·5	30	20	20	–	–	–
3	8·5	30	–	20	–	–	–
4	9·5	30	20	20	–	–	–
5	10·5	30	–	20	37	–	–
6	11·3	30	20	30	28	20	–
7	12·5	–	–	–	50	32	6
8	13·5	–	–	–	37	18	11
9	14·5	–	–	–	–	38	–

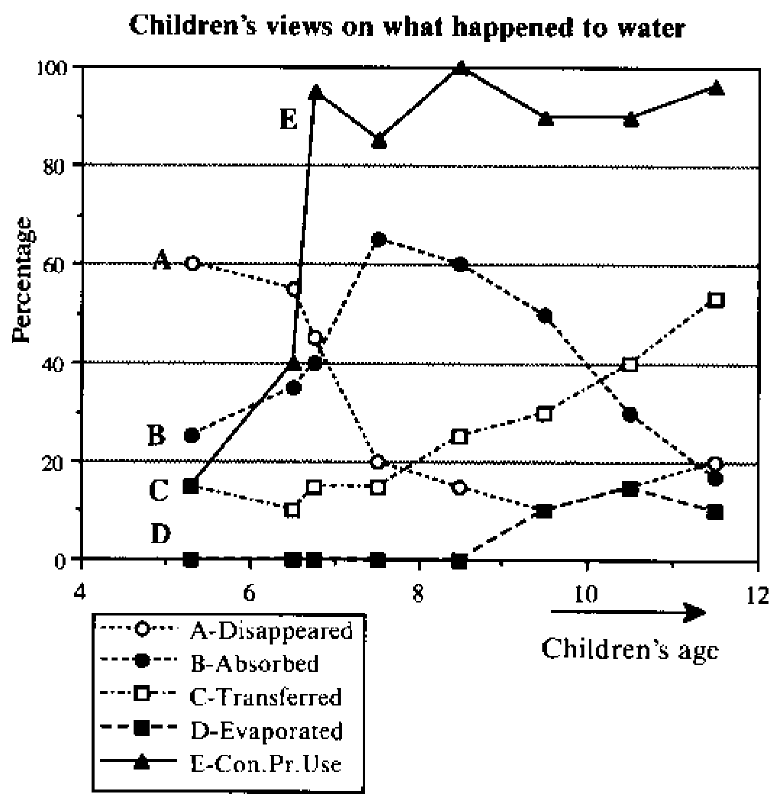
11

The first two phases were just to try and elicit children's thoughts about evaporation. The typical responses were as follows.

- The water disappears or is absorbed into the floor.
- The water has evaporated meaning it is now unseen and being transferred to another medium.
- The water turns into vapor which is made up of small water droplets, mostly unseen, dispersed in air.

- Another explanation related to the last one was that the water had become air.[11]

The following chart helps to describe the responses in phases one and two. A fifth category was added to the chart in order to show the relationship between each age group and their understanding of the conservation of water idea. The general trend is the simpler explanations are more common of younger students and that the more complex notions are typically found in the higher age groups. Also, the older students are more comfortable with the idea of conservation of water and as a result most of their responses more closely mirror what is actually going on during evaporation.



A: 'The water disappeared'; B: 'The water entered into the floor'; C: 'The water is transferred'; D: 'The water evaporated (dispersed)'; E: the use of conservation ideas.

Figure 3. Children's views on what happened to water (floor question) recorded in oral investigation in phase two of the research.

11

With regards to conservation during the study students were placed into 3 separate categories, conservers, non-conservers, and transitional. Those students that were classified as conservers were those who displayed an understanding of the conservation of water for the duration of the interview. The non-conserver group displayed no understanding throughout the interview of conservation. The transitional group students were not consistent with their responses from a scientific point of view. While their logic remained constant, the answers varied between differing understandings of the conservation of water idea.

With regards to the questions about air the students had two responses. They either responded that the air is created by motion or that air is always in existence and that someone or something moves it to create wind. The students who felt that the air always exists were also the same group who responded that water changes into unseen particles and is relocated into a different medium. The vast majority of the students who responded that air is created by motion also said that water had either disappeared or was absorbed by the floor when spilt.[11] This is more evidence showing that students are consistent with their logic about phase change. The results of student's notions of conservation of water were also demonstrated when dealing with air. This shows that the students who answered incorrectly have a consistent misconception with the idea of conservation.

During phase three of the testing students were posed different situations regarding evaporation. The questions were all real world examples as opposed to simply water in a container. This was done to put the device of evaporation into a context and see how the students would respond. One such question dealt with clothes in the laundry.

What was typically found was that given a context, students were able to disregard some of the distracter options during the test. For example, the responses for “the water disappeared” option went down dramatically when students were given context. This also gave rise to the pseudo-scientific answer that “the water changed into hydrogen and oxygen”. Some of the students who came up with that idea responded that the idea just seemed “logical” or “plausible” without really giving a reason as to why their response made sense. The rest of the students that had this view talked in some manner about how hydrogen and oxygen are what make up water as justification for their answer. This was a notion that these students at some point had exposure to in their class work.

What can be derived from each of these studies is that most students have difficulty with matter in the gaseous state. This can be attributed in part to the lack of a visual reference. Without a picture of what’s going on, students have difficulty quantifying what is really going on when matter changes from liquid to gas. Evidence of this can be found in the explanations most students provide when asked to define a gas. As was seen in “Children’s Conceptions of Gas”, students that had the most trouble defining what a gas was simply referenced gases that they had exposure to in their own lives, such as cooking gas. While the gaseous state is one that causes most students difficulty, it is not the only source of misconceptions regarding phase change.

The article “Misconceptualisation of the chemical equilibrium concept as revealed by different evaluation methods” appeared in the European Journal of Science, volume 8, issue 4, in 1986. Written by Gorodetsky and Gussarsky, the article illustrates the difficulties students have with the idea of equilibrium which is important with regards to phase change.

The subjects were all 17-18 years old and had been studying the concept of the chemical equilibrium for a few years in the classes. Their knowledge of the chemical equilibrium was tested by means of two tests. One of the tests was given by the teachers in accordance with the school's curriculum and the teacher's expectations of what the student should know at this point given the depth of coverage the chemical equilibrium subject had in class. The second test was aimed directly at the misconceptions students may have. It was prepared by Gorodetsky and Gussarsky for the purposes of the study. The format was a multiple choice test with 26 questions all aimed at pinpointing different important aspects of the chemical equilibrium concept and what may be causing students to form the misconceptions in the first place. Each question had as an answer both a common misconception and an "I don't know" option.[12]

The results of each test provided the researchers with an adequate way to rank the students and give teachers an idea of how successful their students were at problem solving with regards to the chemical equilibrium. What is more indicative of the students understanding of the subject was the second test written for the purposes of the study. Most students were able to score highly on the tests given to them by their teachers, however the majority of them did not achieve a passing grade of 55% or higher on the misconceptions test. It should be mentioned that the misconceptions test was based on material that all of the students have had ample exposure to in their chemistry classes.

The results were both disappointing and enlightening. They are a very clear example that while the students have learned what has been taught in the classroom, the misconceptions they had previously developed still exist with regards to the main properties of the chemical equilibrium. These results can be used to pinpoint the origins of some of the misconceptions the students have. The most important of these results is that the students again try to use examples from their own experiences in order to explain the phenomenon of chemical equilibrium.[12] This is consistent with what was found in the studies mentioned earlier. Students have a difficult time with abstract concepts and they typically respond by making a connection to something they have seen in their daily lives. That gives them a visualization of what's going on and helps them better explain what is going on. However, this can lead to many misconceptions and poor explanations. The chart below shows the data from the tests broken down by subject matter within the test. For the purposes of this discussion, the individual subjects are not of consequence. What is important to draw from the data is that the student's scores show that they have learned what their teachers taught them, but they still have misconceptions which stem from each student trying to picture what is happening by using real world examples.

Table 3. The percentage of students choosing the correct answer and the misconception distractor.

	Group	*N*	Features A	B	C	D	E	F	G	H
Correct answers	I	88	40	9	42	46	46	53	7	44
	II	74	68	35	61	68	74	81	15	66
	I†	8	88	68	75	100	100	100	38	63
	II†	24	83	79	83	92	96	96	33	92
	I‡	5	100	100	80	100	100	100	40	60
	II‡	21	76	76	81	91	100	100	38	86
Misconceptions	I	88	21	48	41	33	46	32	76	23
	II	74	14	24	31	12	19	15	76	15
	I†	8	—	13	13	—	—	—	68	25
	II†	24	17	4	13		—	4	63	4
	I‡	5	—	20	—		—	—	60	40
	II‡	21	24		14	—	—	—	48	5

† These are subgroups of I and II considered to be good performers on the basis of the number of misconceptions (0–2).
‡ These are subgroups of I and II considered to be high achievers on the basis of the research knowledge test (73–100%).
Note: Knowledge on features **A**, **C**, **D**, **F** and **H** was determined on the basis of 2 correct answers out of 3; for feature **B** it was 3 out of 5; for feature **E** it was 3 out of 4; and for feature **G** it was 2 out of 2.

One thing that should be noted is that this study does not directly pertain to phase change. Phase change is not a form of chemical change; it is a physical change in matter. This means the properties of phase change are in general vastly different from anything that goes on in a chemical change. However, chemical change is often used to describe the state of equilibrium to students in their chemistry classes and that is where some misconceptions in phase change occur.

There have been other articles that have gone more in depth into the struggles students have when learning about the chemical equilibrium. One article in particular is “The Teaching and Learning of the Chemical Equilibrium” written by Van Driel and Graber during 2002 in the journal titled “Chemical education: towards research based practice". The article deals with struggles that student have incurred during other studies on the subject.

When students are first exposed to chemistry, they are given observable reactions to first get a grasp of the central concepts. Chemical reactions are portrayed as proceeding to completion and take place in only one direction, meaning they are irreversible. Completion in this context meaning that the reactants are totally used in the reaction and nothing remains of their original form. A typical example of this would be a combustion reaction where the products when put together burn and what remains is a residue that cannot be salvaged. However in later chemistry classes, students are taught of the reversibility of chemical reactions as well as the possibility that a reaction may not happen to completion. What causes possibly the most problems is the notion that in some cases chemical reactions are really two reactions happening at once in opposite directions. This is better known as Newton's Third Law of Motion, "For every action there is an equal but opposite reaction." This concept causes problems for students because it is nearly impossible to deduce from observation alone.

The inherent difficulties with the observation of chemical equilibrium lead to the formation of many common misconceptions. Students often cannot discriminate between reactions that happen to completion, such as combustion, and those that can be reversed. Often students think that forward reactions happen first before the reverse reaction occurs. The rate at which a reaction takes place and how much of the reactants are used up is also a point of struggle for students. Finally, students often see chemical reactions as speeding up over time until equilibrium is reached, meaning students think that the time a reaction takes is a non-linear function and will accelerate as time passes.[13] All of these notions are issues students have with chemical equilibrium and are difficult to convey for instructors.

Some of these misconceptions can be found across age groups in both secondary and higher forms of education. One that is of particular importance to the idea of phase change is the students view the equilibrium state as dormant, meaning no reaction is taking place and nothing is currently happening.[12] This causes problems understanding what is going on when water is evaporating. Students fail to grasp the cyclical nature of reactions. This causes the students to view a container of water as unchanging over time. However the reality is that evaporation is constantly taking place over time. Without a proper grasp of that concept students continue to respond to questions such as "where did the water in the container go" with simple responses like "it disappeared". Evan the more complex answers such as "the water has been moved to a different medium by the sun" still show a lack of understanding about what is really going on and how the process of evaporation really works.

What the students do to compensate for their own lack of understanding is they memorize facts rather than try to understand the mechanisms of a subject. Graber and Van Driel actually site in their article Gussarsky's and Gorodetsky's work mentioned earlier in this paper. The results are that students will test well in a given class, however when they are questioned on the mechanics of something like phase change or given a test that is riddled with distraction responses pertaining to common misconceptions the students struggle.

From these studies we can see that students benefit greatly from a visualization of the devices they are studying. With a proper visual image, the students will be able to more accurately asses what is truly going on during phase change or any other phenomenon in chemistry. It falls on the shoulders of the instructor to insure that those visualizations are accurately portrayed. Otherwise, students will look to their own experiences to find examples of a subject matter in order to create an image to reference and help explain what is going on.[13] As was shown by Russell, Harlen, and Watt in their study, students will depict the operations of evaporation incorrectly without proper guidance. When asked to draw what was happening during evaporation, most students were unable to accurately draw the process. A phase change activity that can not only accurately show students an image of what is going on, but also allow the students to test their understanding in a lab setting will greatly diminish the commonality of some of the misconceptions regarding phase change.

SCMW Summer Pilot

Purpose

To pilot test The "State Change Microworld" or SCMW, to begin to evaluate learning in this microworld, to test for transfer from Ramp World testing (data has been collected from another study conducted by Professor Janice Gobert), and to gather lessons learned for future development of additional microworlds.

Participants

Participants consisted of 18 middle school students (10 female, 8 male) whose parents volunteered them for our pilot as part of an extended-day offer for summer school programs running earlier in the day at WPI. The students participated in this study with the understanding that their anonymity would be kept. As a result each student will be referred to in this paper by their individual case study number rather than by their names.

Procedure

Our pilot ran for five sessions of two hours each, Monday to Friday. The 18 students were split into three groups over two weeks. The first week had both a 1-3pm group (6 females, 1 male) and a 3-5pm group (4 males) while the second week had only a 1-3pm group (4 females, 3 males).

Each day, students started out with a snack (or lunch for the first 1-3pm group that started eating with the Science ASSISTments Lab at 12:30) before proceeding to a computer lab for the day's "work" on the computers. Time remaining after the computer work would be used for debriefing and socializing. When specific students finished early, they were allowed to surf the web so as not to distract their peers as they finished.

Measurements consisted of a chemistry knowledge pre- test, a scientific inquiry self-efficacy test, and a scientific inquiry skills test developed by the Science Assitments team. These three measures were administered before and after the intervention. Additional instruction in the Control for Variables Strategy was given along with pretest material and CVS posttest was included in post-test material. Finally, a demographics survey was administered with pretest materials only. A general overview of daily activities is provided in Table 1.

Table 1: Summer Science Menu: Student Activities by Day

Day	Activities
Monday	Introductions, Account Setup, Demographic Survey, Chemistry Pretest
Tuesday	Inquiry Self-Efficacy Pretest, Inquiry Skills Pretest "Ramp World" CVS instruction
Wednesday	SCMW Activity (Intervention)
Thursday	Chemistry Posttest, Inquiry Self-Efficacy Posttest, Inquiry Skills Posttest
Friday	Pizza Party, Extra Time (if required), Miscellaneous Amusement

Demographics Survey

The demographics survey included multiple-choice questions regarding SES variables, school grades, and academic interests. Some open-response questions were also included.

Chemistry Knowledge Pre- Test

The chemistry knowledge test included 29 multiple choice questions evaluating general knowledge of middle school science. Many questions were irrelevant to the topic of state change. Those related were often structured around vocabulary and since the intervention did not focus on vocabulary these, too, were unrelated. [Although an area for development, this battery will not likely be used to answer research questions.]

Scientific Inquiry Self-Efficacy Test

The self-efficacy consisted of 14 Likert items regarding students' self-efficacy for their scientific inquiry skills. Thirteen items originated from the science scale of Kettelhut's contextualized self-efficacy battery. Although directly relevant to the pilot, many items included terms that were imprecise leading to some confusion as to what skill students thought they were answering about (although this ambiguity could be resolved with enough background information on their science classrooms).

Scientific Inquiry Skills Test

The skills test consisted of 37 multiple choice items evaluating various dimensions of inquiry skill. [Some] additional open response fields were also included for further explanation.

"Ramp World" (RMW) Apparatus

Interaction with the RMW (i.e. the "gameplay") revolves around manipulating the independent variables and running simulations to test how far a ball rolls down a ramp. Direct instruction explains the CVS concept and questions involve identifying whether or not an initial ramp setup follows CVS.

SCMW Apparatus

The "State Change Microworld" (SCMW) was developed to address the "phase change" related strands of the Massachusetts Frameworks for Natural Science at the middle school level. Very specifically, students should "recognize that a substance (element or compound) has a melting point and a boiling point, both independent of the amount of the sample." As the simulation runs, the user watches a substance melt and then boil as its temperature is being simultaneously graphed as a function of time.

SCMW Activity (Intervention)

The "activity" that makes up the invention exists as a part the ASSISTments System i.e. prompts, input fields, feedback, and anything else other than the flash microworld that is embedded into the system.

Most prompts in the activity asked students to observe the simulation or to manipulate it in a certain way without making reference to any specialized vocabulary (e.g. "independent variable" or "dependent variable").

Student's thoughts were collected via recordings during the intervention. Exam booklets were also distributed for the section of the activity that asked students to design a system for recording data and to use it.

As the beginnings of a cognitive task analysis, we layered the activity to start with very broad prompts, hypothesizing that an expert (i.e. lab members as opposed to even the best students) would give an answer containing all important details and observations while most students would focus in on a few important details while missing others. Because we thought most students would "fail" (at least partially) at the first task, we gradually gave more detailed versions of the same open-text prompts before eventually giving multiple choice questions.

After taking this layering into account, the activity consisted of only two simple parts: 1) observe the "phenomena" being represented by the microworld and note down any important variables; and 2) use the microworld to test what variables effect what's going on (options limited to given drop-downs).

Research Questions

- self-efficacy and content learning
- relationship between independent variables and log files
- transfer of ramp CVS to SCMW CVS
- relationship between inquiry skills and performance
- relationship of self-efficacy to skill as a meta-cognition variable (independent or dependent depending on the context)

Results

For the purposes of this paper what is of most concern is the open response questions the students were asked to fill out. These questions were given ID numbers for ease of use when going over the data and will be referred too in this paper by those numbers. Each student was provided a lab book to use when answering the free response questions provided. Not all students answered every question nor did all the students make use of the lab books provided. The questions were all graded based on the child's understanding over the concepts at work in the given problems. A detailed list of the results can be found in Appendix A.

The first question asked of students was number 53390. The question asked students how they would go about identifying what variables had an effect on the presented experiment. In this example the experiment was a simple flask over a flame with a substance inside of it. The student responses were graded on a 0-3 scale. These scores can be found in Appendix A. The scale was defined as follows:

- A score of 0 meant the student showed no understanding of the subject matter.
- A score of 1 meant the student understood the concept of testing the variables in order to find out which ones affected the experiment. He or she did not mention anything about isolating the variables and did not provide specific examples of the variables they would test.
- A score of 2 meant the student understood the concept of testing the variables in order to find out which ones affected the experiment. He or she understood the concept of isolating a variable to test it properly and gave at least one example of a variable to test.
- A score of 3 meant the student understood the concept of testing the variables in order to find out which ones affected the experiment. He or she understood the concept of isolating a variable to test it properly and gave two or more examples of variables to test.

The responses provided by the students were very polarized for this question. There were five students that scored a 3 on their individual responses. Student Case ID 3 was one of those students and her response was as follows, "To find the answer I would test the different variables at separate times. I would start with the container size, next the amount of substance, the level of heat, and finally the cover. I would try to focus on one variable at a time." The "testing variables at separate times" shows that this student understands both testing variables and isolating them on an individual basis. The student then goes on to provide examples of the different variables that he or she believed affected the outcome of the experiment. This response shows the student has developed the inquiry skills the ASSISTments program is interested in instilling in its users.

Another example of a highly scored response came from Student Case ID 8. This student's response was, "If I were to do this experiment I would use a big beaker so the liquid would have more room to change from liquid to gas or solid to liquid. Also I would use a variation of amounts of water to see if the amount matters in the experiment. Also I would try the experiment about 5 to 10 times with the cap on for half of the tests and the cap off for the other half of the tests." This student chose three variables in particular to manipulate; the size of the beaker, the amount of substance, and the cover on the container. He or she also talks about testing multiple times to see how each variable affects the outcome of the experiment. This student has shown that he or she understands how to think critically to answer the given question.

There was only one student who provided a response that scored a 2. That student was Student Case ID 16 and the response provided was, "I would try first by using different kinds of containers to see if that helps, then I would change the why the blocks of ice were, finally I would see if having the cap on made it faster or slower to melt." This response shows that the student understands which variables are of most concern to the experiment as well as that each one needs to be tested. What this response does not mention specifically is that the variables need to be tested separately from each other. While the subject knows that there are variables that need to be examined, he or she does not show an understanding that to properly determine its effects each variable must be tested individually. Without this specified in the response it is unclear if the student understands the concept of isolating each variable.

There were five students that provided responses that scored a 1. One such student was Student Case ID 13 and the response from the lab book was, “If you wanted to find out if the container size, amount of substance, and whether the top of the flask is capped affect what happens inside the flask then you would just to the experiment over with different variation and record what you did.” This response does show that the student understands the need to test the variables that are at work; however he or she fails to go into any detail as to which variables are the ones that should be focused on. Additionally, the student does not mention that the variables need to be tested on an individual basis. From the response it would appear the student is talking about testing all the variables at once and seeing what happens.

Another example of a response that scored a 1 came from Student Case ID 14. “I would take different size containers of the same kind and test them with different sized ice.” This response does mentions testing some of the variables at play; however the student does not mention anything about isolating the variables at play in the experiment. This response, as well as the response of Student Case ID 13 shows a lack of inquiry skills that the ASSISTments system will aim to target to help this student better understand the larger concepts.

During the testing there were two students that received a score of 0 on their responses. In both cases the students’ responses had nothing to do with the experiment. These responses did not detail any variables or talk in anyway about testing anything specific. There were also three students that did not provide an answer.

The next open response question that was noteworthy came from question 53392 part C. The question dealt with the exact same example as did question 53390 and was graded on a similar 0-3 scale. This time students were asked specifically if the container size had anything to do with the results of their experiment. The goal was to get the students to draw a conclusion given the data that they collected when running the experiment. On the whole students typically did not answer this question. Only six students provided responses and the range of answers went from 0-3. The full list of scores can be found in Appendix A. The grading system was as follows:

- A score of 0 meant the student showed no understanding of the subject matter.
- A score of 1 meant the student was able to draw a conclusion based on the data he or she collected. There was no evidence taken from the data provided in the response.
- A score of 2 meant the student was able to draw a conclusion based on the data he or she collected. The response also included some evidence drawn from the data each student collected to support the answer. There was either no evidence provided that the student understood the concept of the problem or the conclusion he or she drew was incorrect.
- A score of 3 meant the student was able to draw a conclusion based on the data he or she collected. The response also included several pieces of evidence drawn from the data each student collected to support the answer. There also was evidence that the student understood the overall concept of the problem at hand.

There was one example of a response that scored a 3 and it came from Student Case ID 2. The response was, "When I did the sim with the small flask it took 48 min to go through all 3 stages. When I had a medium flask, the cycle completes in 42 min. When I had the big flask it took 38 min for the cycle to complete. So it is apparent through my data that the flask size changed something." This response shows that the student looked over the data that he or she collected, identified a trend, made a conclusion, and used the data to justify the conclusions he or she made. This response shows the type of thinking that the ASSISTments program aims to show its users how to implement.

Two students provided responses that were given a score of 2. One of those students was Student Case ID 12 who wrote, "I believe it affected what happened because the bigger the flask the easier and longer it takes you to melt and heat the liquid. I think thinks this happens because the more area there is in the container the more you have to heat up." This response illustrates a clear conclusion that the student has drawn from conducting the experiment in the micro-world. The problem with this response is that the student has drawn an invalid conclusion for his or her own data. This stems from a simple misconception that the ASSISTments system aims to solve. The student understands how to go about testing things, but does not yet understand what exactly is going on.

The other student who gave a response that received a score of a 2 was Student Case ID 5. The response was as follows, "Well when it is smaller there is less room so the ice melts faster because there's less room so it has less time to not melt." This response shows that the student manipulated the container size in his or her experiment and drew the conclusion that the smaller container made a difference in the time elapsed. The issue here is specific data points were not mentioned nor were they specific to talk about leaving the other variables constant while testing the container size. This attention to detail is something the ASSISTments system will aim to show its users to incorporate.

There were two students whose responses earned them a 1. Student Case ID 8 provided one such response, “I think the container size affected of how quickly the solid melted.” This response shows that the student did in fact draw a conclusion; however he or she fails to mention how this conclusion was reached or provide data supporting the theory. In terms of a scientific response this is not a valid conclusion. For something to be stated as fact there needs to be evidence to back it up. This student does not show an understanding of this concept and the goal is that the science ASSISTments system will show this subject how to properly defend what it is that he or she states with supporting facts. As a side note there was also one student who responded and received a 0 while the rest of the students skipped this question.

The next question of importance was number 53393 where the students were asked about the amount of substance and whether of not it had an affect on how quickly time elapsed during the experiment. The question dealt with the exact same example as did question 53390 and was graded on a similar 0-3 scale. The goal was to get the students to draw a conclusion given the data that they collected when running the experiment. The students typically scored low to moderate on this question. Eleven of the students provided responses and the range of answers went from 0-2. Appendix A contains the full list of scores. The grading system was as follows:

- A score of 0 meant the student showed no understanding of the subject matter.
- A score of 1 meant the student was able to draw a conclusion based on the data he or she collected. There was no evidence taken from the data provided in the response.

- A score of 2 meant the student was able to draw a conclusion based on the data he or she collected. The response also included some evidence drawn from the data each student collected to support the answer. There was either no evidence provided that the student understood the concept of the problem or the conclusion he or she drew was incorrect.
- A score of 3 meant the student was able to draw a conclusion based on the data he or she collected. The response also included several pieces of evidence drawn from the data each student collected to support the answer. There also was evidence that the student understood the overall concept of the problem at hand.

None of the students got the full 3 points possible with this question. Three of the students did manage to score a 2 on their responses. One such student was Student Case ID 12. The response was as follows, "Yes because the more substance the longer it takes because the more there is to melt." This response shows a clear understanding of the concept that more substance will take longer to melt than a smaller amount. This student did not however provide specific examples of data to reference to and defend his or her conclusion making it hard to conclude if the student has done the experiment properly and isolated variables while testing the amount of the substance.

The largest group of students received a 1 on there responses. Five students were in this group, among them Student Case ID 2. The response read, "The more substance, the longer it takes for the cycle to run through. It took less time for the small substance to melt than the large substance." This response shows a conclusion that has clearly been drawn by working in the micro-world and manipulating the variables. The student failed to show that he or she understands how to isolate just the amount of substance when testing it nor does he or she provide any hard data to support the claim.

Another example of a response that scored a 1 came from Student Case ID 17. “I think so because a smaller piece melts faster.” A conclusion here has been made but the student does not back it up with any data nor does he or she explain why the claim was made at all. This makes it impossible for the corrector to determine if the subject understands the larger concepts at work in each problem. The goal of the ASSISTments system will be to show these students how to verify and support their claims with data and to teach the students how to use inquiry skills in their own reasoning. Three of the students scored a 0 on their responses while five simply skipped the question.

The final open response question was number 53394 part C. The question dealt with the exact same example as did question 53390 and was graded on a similar 0-3 scale. The goal was to get the students to draw a conclusion about the cover on the beaker given the data that they collected when running the experiment. The students on the whole skipped this question. Only four of the students provided responses and the range of answers went from 0-2. Appendix A also has the full list of scores. The grading system was as follows:

- A score of 0 meant the student showed no understanding of the subject matter.
- A score of 1 meant the student was able to draw a conclusion based on the data he or she collected. There was no evidence taken from the data provided in the response.
- A score of 2 meant the student was able to draw a conclusion based on the data he or she collected. The response also included some evidence drawn from the data

each student collected to support the answer. There was either no evidence provided that the student understood the concept of the problem or the conclusion he or she drew was incorrect.

- A score of 3 meant the student was able to draw a conclusion based on the data he or she collected. The response also included several pieces of evidence drawn from the data each student collected to support the answer. There also was evidence that the student understood the overall concept of the problem at hand.

There were no responses that received a grade of a 3.

Two of the students did provide responses that met the requirements for a grade of 2. Student Case ID 5 was one of those students and the response provided was, “The stopper did affect it because it didn’t let anything out the flank so the heat stays in and melts faster.” This response shows that the student manipulated the variable in question, in this case the stopper on top of the beaker. The response also draws a valid conclusion from the student’s experimentation in the micro-world. What this response lacks is any real data to support the claim. Without the data, as has been said several times already, it becomes hard to tell if the student isolated the variable and therefore brings his or her logic into question. By providing data to support an argument, the student shows they do not have the misconceptions that are typically present when dealing with the subject matter. If those misconceptions exist, it is the goal of this project that the ASSISTments system will be able to flush out the issue and make it easier for teachers to identify what students need help and where they are struggling.

Student Case ID 12 was the only student to provide a response that received a 1. It reads, “It affected the temperature and how high it was in the end of the experiment.”

This response does illustrate a conclusion that was drawn; however there seems to be no reasoning for it. There is no data to back up the claim nor does the claim itself appear valid. It is unclear why the student came to an invalid conclusion because there is no data or reasoning provided in the open response answer. There was one other student who scored a 0 and the remaining students simply left the question blank.

It should be noted that these students were middle school or under in terms of grade level and age. The questions that were skipped may not have been skipped from a lack of understanding but more from a sense of fatigue that set in after completing other questions in the test. This could also explain some of the open response answers that were too brief and therefore scored very low overall.

Conclusion

The Science ASSISTments system was successful in tutoring this small group of middle school students' inquiry skills in the domain of state change. For each question the individual students had to provide an explanation for the different scenarios presented to them after they had a chance to manipulate the micro world. The higher scoring responses for each question met the requirements set forth, namely to draw a conclusion based on the data he or she collected; include several pieces of evidence drawn from the data, and an overall understanding of the problem at hand.

The high scoring student responses hit on all the major points. Some note-worthy examples read, "Also I would try the experiment about 5 to 10 times with the cap on for half of the tests and the cap off for the other half of the tests." Students spent time going into greater detail about the decisions they made when testing and the conclusions they drew. These responses show a clear method that each student followed as well as the ability to isolate what is really going on in a given situation.

The benefits of the Science ASSISTments system can be even better portrayed by the lower scoring responses. For example a score of 1 meant the student was able to draw a conclusion based on the data he or she collected. While there was no evidence taken from the data provided in the response, the student in question still was able to draw conclusions based upon manipulation and visualization of the micro world. When reading these responses you are able to find a clear process that the student is going through. In short, the Science ASSISTments system should be tested on a larger classroom of students to illustrate the full benefits it would provide and to get a better sample size for data collection. Given the sample size of the current data, the ASSISTments system is an effective way to tutor middle school students' inquiry skills in the domain of state change.

Appendix A

Case ID	Gender (1=male 2=female)	TestingSession	Lab book	chemPre	scChemPre	cvsPre
5	2	3		6	2	2
3	2	1		15	3	4
8	1	3	lab book	15	1	1
9	1	3	minimal	19	6	6
15	2	1		9	3	4
17	2	3	no lab book	15	5	2
4	2	1	lab book	14	6	3
11	1	2		17	4	6
13	1	1		18	4	3
16	1	2	lab book	16	5	3
14	2	1		15	2	3
2	1	2		12	3	4
6	2	1	has lab book	16	6	6
1	2	3		17	4	3
12	2	1	lab book	12	3	2
18	1	3	lab book	?	5	1

Case ID	IVPretest	DVPretest	53390	53392 evidence	5339c2 rea	53393 c1	53393 c2	53394 c1	53394 c2
5	2	1	NA	2	1	0	0	2	1
3	3	3	3	NA	NA	0	0	NA	NA
8	3	1	3	1	0	1	0	NA	NA
9	1	2	1	NA	NA	NA	NA	NA	NA
15	4	5	1	NA	NA	1	0	NA	NA
17	0	3	0	NA	NA	1	1	NA	NA
4	2	5	3	NA	NA	NA	NA	NA	NA
11	3	5	NA	NA	NA	1	1	NA	NA
13	3	2	1	NA	NA	2	1	NA	NA
16	0	2	2	NA	NA	NA		NA	NA
14	2	1	1	1	0	1	0	2	1
2	4	2	3	3	2	2	1	NA	NA
6	6	6	3	NA	NA	NA	NA	NA	NA
1	0	1	NA	NA	NA	NA	NA	NA	NA
12	2	0	0	2	1	2	1	1	0
18	2	4	1	0	0	0	0	0	0

References

1) *Assistments*. Web. Mar.-Apr. 2010. <http://www.assistments.org/>.

2) *MAC - Teaching High School Science through Model-based Technology*. Web. Mar.-Apr. 2010. <http://mac.concord.org/>.

3) Gobert, J., Heffernan, N., Ruiz, C., & Kim, R., (2008). Annual report for AMI: ASSISTments Meets Inquiry (NSF-DRL# 0733286). Submitted to the National Science Foundation.

4) *Smithsonian Institution*. Web. Mar.-Apr. 2010. <http://www.si.edu/>.

5) "Science and Technology Concepts for Middle Schools (STC/MS)." *STC/MS - Introduction*. Web. Feb.-Mar. 2010. <http://www.stcms.si.edu/stcms.htm>.

6) "A Brief Overview of The National Science Resources Center." *National Science Resources Center (NSRC) Home Page*. Web. Mar.-Apr. 2010. <http://www.nsrconline.org/>.

7) "Information about the National Science Resources Center." *National Science Resources Center (NSRC) Home Page*. Web. Mar.-Apr. 2010. <http://www.nsrconline.org/about_the_nsrc/index.html>.

8) "How STC/MS™ Addresses the National Science Education Standards." *2.0 Home*. Web. Feb.-Mar. 2010. <http://www.carolinacurriculum.com/>.

9) *National Science Resources Center (NSRC) Home Page*. Web. Mar.-Apr. 2010. <http://www.nsrconline.org/curriculum_resources/middle_school.html#ResearchDev>.

10) *The International Journal of Science Education*. 5th ed. Vol. 10. 1988. Print.

11) *The International Journal of Science Education*. Special ed. Vol. 11. 1989. Print.

12) *The International Journal of Science Education*. 2nd ed. Vol. 16. 1994. Print.

13) *European Journal of Science*. 4th ed. Vol. 8. Print.