Masterminding Science Education:

Cost and Pedagogical Potency in Science Education

by

Paul A. Wagner
University of Houston - Clear Lake
and
Lillian Benavente-McEnery
University of Houston - Clear Lake

There is no shortage of criticism about education in general and science education in particular. In the wake of such criticism, there are many recommendations detailing how to get science education back on the right track (Nisbet, 1993; Stenhouse, 1985; Venville, Adey, Larkin, 2003; Yager, 2000). Frequently, these proposals entail the adoption of costly curricular packages and instructional technologies. Whatever the merit of the individual proposals, each generally requires a considerable financial investment on the part of adopting institutions. Laboratory equipment and computers are expensive, but must every solution require such extensive investment?

The Infrastructure of Experimental Practice and the *Feel* of Science

While we argue for the importance of initiating children to experimental practice as early and as often as possible (Wagner, 1983; Hammrich, 1997; Yager, 2000), our principle area of concern continues to evolve around techniques for introducing children to the conceptual foundations of science (Wagner, 1986; Frykholm, 2005; Gallagher, 2000). In practice, while the preponderance of scientific effort swirls around experimental achievements, conceptual achievements continue to be astoundingly important in the overall advancement of science. For example, in the twentieth century, Albert Einstein, Richard Feynman and S. W. Hawking are recognized as scientists extraordinaire, yet none of them spent much time in laboratories. Of course, without the empirical data uncovered by the great experimentalists, such as Michelson and Morley and most winners of the Nobel in physics, theoreticians would have nothing to go on. Science requires a symbiosis between conceptualist and experimentalist alike. The dynamic relation between theory and experiment must be made conspicuously clear to students as early as possible if they are to have any hope of understanding a "scientific spirit or gestalt." That is to say, a worldview animated by a spirit of bold inquiry and moderated by constraints of logic and empirical data. In fact, the National Science Education Standards (1996) state, "All students should develop abilities necessary to *do* scientific inquiry and develop an understanding *about* scientific inquiry."(Italics added.) The American Association for the Advancement of Science (1993) affirmed this worldview concluding, "Science instruction should facilitate the development or understanding of science as a way of knowing."

Science progresses as much from an elegant thought experiment as from a laboratory experiment (Cohen, 1985; Hug, Krajcik, & Marx, 2005; Miller, 1984; Aris, Davis, & Stuewer, 1983). Science has, on more than a few occasions, advanced as a result of a scientist's elegant theoretical deductions as much as it has from the introduction of a new research instrument (Ackermann, 1985; Harre', 1983). Both experimental imagination and new technical constructs of mind and laboratory coalesce to bring about the ever-growing success of science. For science education to be successful, the sources of all that animates scientific achievement should be made evident to students as early as possible. The importance of laboratory technique and focus on nature's responsiveness to human inquiry remain central to all science education. But so, too, should the conceptual manipulation of experimental results. Such manipulation often takes place in the language of equations, but equally central is the general language of logic. The challenge is how to get all this into the minds of innocents not fully at grips with the impending onslaught of even their own puberty. What must be the price: both economically and in terms of "time on task"?

The Game of a Deductively Organized Science

In what follows, we explain how an inexpensive game is used to introduce students to the cognitively active features of scientific research. These features include hypothesis formation, deduction, and responsiveness to empirical results. Of course, as often gets misunderstood, it is not the game that teaches children, but what a teacher *says* before, during and after a game that makes it pedagogically useful.

The game Mastermind can be used to teach students about several key concepts in science, including the notion of deduction, closed system, good test, logic as a language of science, the purpose of an experiment, elegance of expression, the role of thought experiments, and methodological values and practices. In addition, there is value in learning from professional publication, the importance of meticulous record-keeping, the function of a research team, and the function of a research-team leader. Carefully managed discussion exploits the interest children have in game-playing and introduces them to the conceptual foundations of deductively-organized science.

The game Mastermind can both motivate students and illustrate for them the conceptual characteristics of scientific engagement. The game is a powerful and inexpensive tool for teaching about deductively organized sciences. Its potency, however, can only be realized through the skillful pedagogy of the instructor (Adams & Krockover, 1997; Frykholm, 2005; Haney, Czemiak, Lumpe, 2003).

Good teachers make learning happen far more readily than any curricular tool alone. Consequently, in terms of cost-benefit analysis, research and service dollars are better spent in further and better training for teachers than on expensive capital equipment. That is the motivating impulse of practices initiated at the Institute for Logic and Cognitive Studies at the University of Houston- Clear Lake. The specifics of the discussion that follow are based on actual teacher-training practices and experimental classes for children grades four through ten held at the Institute. The Institute is situated near the Johnson Space Center and many students have parents who are engineers and space scientists. These parents often express amazement that their children have a much greater *sense of science* after a week at the Institute than after even years in school. The students may know that Mom designs space technology and that Dad interprets tracings on photographic plates, but it is quite another thing altogether to understand science as do practicing engineers and scientists.

Knowing what scientists work on is entirely different from *feeling* the success of powerful and meaningful deductions (AbuSharbain, 2002). The Institute's use of a simple game of Mastermind has ostensibly done more than previous school curricula or even sophisticated parental involvement. The able use of Mastermind takes students beyond "mere" observation and semantic mapping of concepts to a *sense of* scientific discovery.

Scientific discovery is rarely precipitated by serendipity. Rather, it comes on the heels of experimental planning and confrontation followed by subsequent accommodation of theory.

The game Mastermind can be purchased in nearly any toy store for the modest sum of ten to fifteen dollars. This game has been available for twenty or more years. Its potential as a curricular resource is often overlooked in favor of expensive and well-marketed gadgetry. Yet, with appropriately trained teachers, Mastermind can serve as a nearly invaluable resource. At the very least what follows will show that advancement in science education, specifically but probably all education, need not require additional expensive gadgetry. For example, there is a software version of Mastermind available, but it does no more pedagogically speaking than the decades old game with plastic game board and pegs.

Play of the Game

The point of Mastermind is to figure out a secret code in as few steps as possible. The code is created by an ordered set of colored pegs.[1] The game consists of a playing board with ten rows. Each row contains four large holes in which large colored pegs may be placed. There are six different colors of pegs from which to choose. Adjacent to each row, there is a 2 x 2 matrix of small holes. The matrix of small holes is used to provide information in response to each attempt at de-coding and the information is conveyed in each matrix by placing small black or white pegs indicating one of three truths for each peg placed. The matrix conveys data as follows: First, a small black peg signifies that a correctly colored peg is in the correct hole in the particular row being tested. The black peg does not reveal which hole in the row, or which color is correct. These latter two bits of information remain to be figured out subsequently by the code breaker. Second, each small white peg in the matrix indicates that there is a correctly colored peg in an improper hole, again, in the row under test. Third, each hole without a peg indicates the remaining large colored pegs not at all belonging to the coded sequence.

As the game is played at the Institute, each player trying to uncover the secret code is identified as a scientist. The code maker is referred to as "Nature." Each row of large colored pegs proffered for the test is described by teachers advising students on the play of the game as a *hypothesis*. Each subsequent piece of data revealed in the matrix is referred to as a *test*. The hypothesis and the test together are referred to as an *experiment*. To set up play, all scientists look away as "Nature" creates an ordered sequence of four colored pegs secreted behind a shield on the playing board.

Play begins when a scientist (or a team of scientists) places four colored pegs in the first row of the playing board; call this code, nature's *law*(NL). Nature responds, that is to say, completes a *test* of the young scientist's hypothesis, by placing small black and white pegs in the accompanying four celled matrix to the left on the playing board in the manner described above. Nature's response tells the scientist (or scientific team) how closely she (or they) have succeeded in approximating NL. Consider play in a typical game wherein one might find the following hypothesis to be tested:

---

[1] The most recent commercial version of the game uses red feedback pegs rather than black ones, and contains orange instead of Black. This paper incorporates the "classic" set of color names: Black, White, Red, Green, Blue, and Yellow, and refers to feedback as "Black" or "White."

NL Locations A, B, C and D refer to the four holes on the playing board from nature's left to right. The color words refer to colored pegs.

Location A	Location B	Location C	Location D
Black	Red	Green	Red

(**H1**) The first hypothesis to be tested is as follows.

Green	Yellow	Yellow	Blue

The first hypothesis to be tested is called a eutochia at the Institute. Eutochia comes from the ancient Greek; Aristotle used it to mean a reasoned guess (Lesher, 1978). Since nothing more is known about nature's law than the constraints of the game, this hypothesis is distinguished from subsequent hypotheses as being based solely on the arbitrary constraints defining the realm of investigation. The eutochia represents no accumulated knowledge resulting from experimentation. Accumulated knowledge begins with the test results that complete the experiment. In the unfolding example, the completion of the first test results is outlined:

(**T1**) Nature responds to test by revealing the following information in the matrix:

1 white peg

The placement of the peg in one of the four small matrix cells addresses color only. It also means that the correctly colored peg is in the wrong location. The teacher now has the students list each bit of information revealed by T1. (The symbolic representation in the brackets is usually added later as will be explained below.)

1. If green(G) is in nature's law(NL), then blue(BL) and yellow(Y) are not, [G-> - (BL v Y)]

2. If blue is in nature's law, then green and yellow are not, [BL-> -(G v Y)]

3. If yellow is in nature's law, then blue and green are not, [Y-> -(BL v G)]

4. If green is in nature's law, then it is not in the first location counting from the left, [G-> -G/A]

5. If blue is in nature's law, then it is not in the fourth location counting from the left, [BL-> - BL/D]

6. If yellow is in nature's law, then it is not in the second or third location counting from the left, [Y-> -/(B or C)]

7. If yellow is in nature's law, then there is only one yellow peg in the code [Y<->1]

8. If green is in nature's law, then there <u>may be</u> more than one green peg [G->G may be greater than 1]

9. If blue is in nature's law, then there <u>may be</u> more than one blue peg. [Bl->Bl may be greater than 1]

The teacher continues asking students about what they know from the test until all the information recoverable from a given test is gathered. This takes extraordinary teacher talent, but this is what helps students understand the meticulous and logical path typical of most real discovery in the deductively organized sciences. Each scientist or team playing should have their identification of universally recognized knowledge identified in the published results which will take place later in the process. This identification will become important in subsequent post-game discussion about the nature of scientific contribution and the effort to evaluate some contributions as somehow more worthy than others.

When students have exhausted all that is known as a consequence of T1, it is then time to begin a second experiment. This begins with framing H2. For purposes of illustration, imagine the students have done this as shown below.

H2	Black	Red	Yellow	Green
T2	Two black pegs and one white peg			

Nature's response to H2 confirms that two large colored pegs in H2 match both the color and location of same in NL. T2 also shows that a third peg is the right color but located improperly. Once again the teacher asks the student scientists to list exhaustively all that they know. This list of test results is even longer than the list completed at the end of the first experiment. The published test results of T2 must reflect the results of T2 in light of accumulated results from T1. Journal space forbids listing the results of T2 and for purposes of future discussion, the listed results of T1 will suffice. For purposes of illustration, imagine T2's published results are in place, and students proceed to create the most *powerful* test possible for the third experiment. For simplicity sake, we offer the H3 for a subsequent test.

H3 Red Green Red Black

T3 Four small white pegs

Four white pegs show that all colored large pegs in NL have been identified and that none of the hypothesized locations is predictive. Once again the teacher lists the published results on the board before allowing the students to begin work on H4. Again, it is left to the reader to work out the results of T3. The careful reader doing so will recognize the excitement of the more able students who part way through the listing of T3's results will recognize that they now *know* NL and H4 will just be a matter of confirming that truth.

H4 Black Red Green Red

NL confirmed

Because of journal space limitations, the above game is short and reflects an uncommon fortuitousness in the third test of the scientist. Nevertheless, this example illustrates how each step of the game can be studied as a conceptual routine common in any deductively organized science such as physics or molecular biology. In what follows we will now discuss the scientific principles illustrated throughout the game. The discussion will begin with an overview of the purpose of science and experiment. It will then proceed to discuss the following concepts: closed sciences and deduction, eutochia, record keeping and publication, thought experiments and the good test, elegance and the language of science, evaluating scientific contributions, serendipity, research teams, and finally refutation and confirmation.

The Purpose of Science and Experiment

There are a variety of spins describing what goes on in science. Some of these have leaked into the practices of science education as well. Perhaps the most popular spin portrays science as reading off a litany of nature's truths through a process of deft observation. That notion is not only extraordinarily naïve, but also glosses over the difficulty of coming to grips with the notion of truth itself. Science can never guarantee the arrival of truth. At its best it can assure researchers when they are moving ever further from error and toward greater control over the world around us. The search for truth is a worthy ideal for scientists to possess, and certainly there is no harm in the earnest seeking of truth as long as scientists keep in mind human fallibility and consequent tentativeness that often accompanies the grasp at suspected truth. The notion of truth is slippery to the point of being, at times, unwieldy (Lynch, 2004 :Horwich, 1990 : Williams, 2002: Blackburn, 2005: Dummet, 2004; Davidson, 2001, 2005: Putnam, 1975). However, when one goes about defining truth, it is truth, in at least some sense, at which scientists aim. This is the case even if the most optimistic can never confirm convergence on truth writ large (Peirce,1905).

Most scientists, then, agree that science is self-correcting and moves ever further away from error. This notion of an alleged self-correcting drift in science (Ayer, 1968; Rescher, 1978), risks creating expectations that certainty might be reached, however, this expectation must forever be fought against.

Science aims at truth but settles for the best possible explanation. As might be expected, there is much debate about what counts as the best possible explanation in science. Karl Popper, perhaps the best known philosopher of physics for most of the twentieth century, thought the best explanation was one which survived rigorous tests to falsify it. Popper described such resiliency as verisimilitude. Verisimilitude gave an explanation of the best possible betting odds until something better came along. It is, in short, the closest science can come to truth. Too, Popper was convinced that while we can never know an affirmation for certain, it is possible to know for certain that a deduced theory or experimental hypothesis has been refuted (Popper 1958,1972). For example, the corroborated observation of one green duck falsifies with absolute certainty a hypothesis that all ducks are white. But, inasmuch as the devil lurks in the details, commentators have pointed out that in real science, refutations are sometimes ignored because a theory is so treasured that the semantics or observational interpretation may be re-arranged to protect the hypothesis from an otherwise air-tight refutation produced with ostensively deductive certainty. Popper's thinking remains heuristically useful as will be shown a bit later, but for the time being it must be acknowledged that his ambition for certainty of a sort is too ambitious. In short, there is no crucis experimentum anywhere in real science – not of the confirmatory sort Newton hoped for and not of the sort Popper proposed (Popper, 1963).

The dream of a crucis experimentum (Zahar, 1978), wherein truth about natural law is definitively obtained or dismissed once and for all, is a myth. Science is not a matter of shooting bullets at a target called truth. Rather, science is a matter of creating research programmes comprised of heuristical strategies for moving researchers away from error and presumably towards truth (Lakatos, 1978). In sciences such as physics and molecular biology, these heuristical strategies work to preserve the rigor of deductive analysis as Popper proposed, but refutations are now seen as discrediting bits and pieces of a hypothesis or theory and not the research programme generally (Papineau, 1979).

Physicists and other deductively predisposed scientists begin a research programme by stipulating the boundaries of a closed system. For example, when working on the general theory of relativity, Albert Einstein stipulated (among other things) that nothing can travel faster than the speed of light. And, as a hundred years of additional research have revealed, Einstein's top down system seems again and again to be on the right path. Even so, the road has been a bit bumpy at times. If too much emphasis had been placed on either refutation or confirmation of Einstein's deductions, scientists could have lost one of the most powerful research programmes ever devised.

Science is hard, *reflective* work (Gallagher, 2000) and not merely a matter of poking and prodding nature to collect truth. Consider, for example, a very famous "near miss" at approximate truth wherein deductions inferred from relativity theory seemed incommensurable with an early set of observations. Sir Arthur Eddington traveled to the Cape of Good Hope to photograph a solar eclipse and thereby confirmed deduced predictions from Einstein's theory of general relativity. Eddington took three photographs; unfortunately, only one of them confirmed general relativity. With a score of two against confirmation and one in favor, it would seem the theory went unconfirmed. Moreover, add to such a result the thinking of Popper and his disciples, the first refutation should have tanked the theory altogether right from the start. But Eddington persisted and when the third photograph confirmed relativity, Eddington dismissed the results of the first two photographs as a product of human error - very convenient error indeed. Nevertheless, the decision to write the disconfirmations off as human error proved enormously valuable to science in the long run. Interestingly, it was not until thirty years later that unequivocal, physical confirmation of general relativity theory was found. For thirty years, an uncertain theory served physics by implying deductions that were tested and led more deeply toward nature's secrets. Einstein's own characteristic comment to journalists on the report of Eddington's alleged confirmation of general relativity was prophetic. Einstein said he would be sorry for God if He had failed to recognize the elegance (deductive rigor) of the theory.

Philosophers of science today now believe that especially in physics the scientist's immediate goal is to detect deviance from a hypothesized pattern. Any deviance is then addressed either through further disconfirming experience or accommodated through modification of originating principles. If anomalous information fails to lead to scientific advance, it is dismissed as trivial or, more likely, errant (Achinstein, 1985; De Boer, 1991; Oliver & Nichols, 2001). In short, the spin that science simply reads a litany of nature's truths will not wash. Science is about doing further science! (This is not to deny that engineers and others do remarkable things with the findings of science as humans extend their control over the world- they do. But the scientist qua scientist is about doing science and not just doing something with the findings of science.)

There is an interesting example of science's focus on doing science that is especially instructive here. The example took place in the mid-1980's and involved a young graduate student by the name of Mark Spector. Spector believed his dissertation advisor's controversial theory about cancer was on the mark (no pun intended, though it is funny in an ironic sort of way). Spector set out to demonstrate the "truth" of his advisor's theory. Unable to find evidence supporting his advisor's premonitions, Spector faked an experiment. By treating suspect monkey cells with iodine following an investigative procedure (Wade & Broad, 1984), Spector falsely got confirmation of his advisor's theory. His research was immediately hailed as a major breakthrough in cancer research. Unfortunately, as the young graduate student discovered, science isn't about listing truths. It is, as noted above, about adding to, deleting from or otherwise amending a current research programme. In the most exotic cases, it *may* lead to a paradigm shift but only to accommodate anomalies accumulated over the years in concert with most of what is effectively described by any current research programmes.

As other researchers tried to replicate Spector's experiment to get work product for further investigation, each failed. Without work product there was no hope of advancing the emerging new paradigm. Eventually, Spector was observed by jealous colleagues compromising his data with iodine. Spector admitted his fraud and the episode came to an end, as did Spector's hope for a career in science. Such scandals teach important lessons about science. Science is simply not about stabbing at truths. It is, in large part, about creating work product that can subsequently used by other scientists to draw the community of scientists ever further away from error.

The research strategy employed in current deductively organized sciences is usually described as the hypothetical-deductive method. While this method has shortcomings and doesn't apply as broadly to scientific research as once thought (Van Fraassen, 1980; Suppe, 1978), it continues to prove of enormous value in physics, molecular biology, artificial intelligence research and so on.

In using Mastermind to teach students about the deductively organized sciences, students should be asked to describe how Mastermind differs a bit even from those sciences. The most notable difference to get across is the fact that in Mastermind there is always a final and certain conclusion. In science, however, even in physics, the most securely deduced conclusions can only be tentatively adhered to. Finally, although there is no way in which Mastermind gives students a thorough sense of the observational, semantic mapping and economic dynamics of actual science, it can give students a sense of the crucial role of the logical apparatus necessary to scientific thinking. That is no small achievement (Haney, et al, 2003).

Closed Sciences and Deduction

As noted above, deductively organized sciences begin organizing by creating closed systems. In Mastermind, similarly, the closed system of concern is limited. Specifically, in Mastermind, it is limited to six differently colored pegs (data) and four holes- "conceptual (semantic) spaces" to be used for organizing data into a pattern that eventually mirrors NL. Even the most naïve student intuitively recognizes that there is a limited number of possibilities (1296 to be exact) for NL.

Whether in science or Mastermind, systems are closed not by Nature, but by the postulates of scientists. For example, Einstein's postulate that nothing travels faster than the speed of light allowed for deductions that could not otherwise have been made. Yet not only Einstein, but those that followed, knew that such postulates do not limit nature but are of heuristic value (d'Espagnat, 1983; Frank, 1947; Paris, 1982). For example, if something did travel faster than light, there was no way to study it (Toulmin, 1958). Bounded frameworks for deduction allow inferences for testing and nothing more. This testing is, in the end, what makes science an empirical study (Watkins, 1984; Tarski, 1956; Salmon, 1967; Scheffler, 1982; Schlick, 1925).

Deductions are truth-preserving techniques of inference-making. They let scientists extend their understanding from well-entrenched commitments to hypotheses worthy of testing (Goodman, 1978; Holloway, 2000). As a matter of formal logic, any conclusions deduced are true--given the truth of the premises. If empirically tested conclusions prove wayward in a deductively organized science, scientists learn there is a problem with their premises or presumed background knowledge.

The analogy of science to Mastermind makes it easy for students to understand why scientists create closed systems to identify testable hypotheses. For example, by eliminating one of six colors in NL, the range of possible matches is reduced by nearly one half. Students see in such a reduction an extraordinary advance in knowledge.

Again, unlike real science, in Mastermind there are no test results that can throw doubt upon the *defining postulates*. This is because the game is no more than a set of rules and definitions. Unlike closed systems utilized in actual scientific practice, the game Mastermind "owes no debt" to an outside world to which it is forced to be responsive. Nonetheless, the apparent inferential architecture of deductive science and Mastermind are identical.

Even within a modified closed system, such as the children's game Twenty Questions, enormous amounts of information are acquired from few tests (questions responded to), if the questions are carefully constructed. Specifically, as Nicholas Rescher (1978) points out: "One knows from information theory that if the game Twenty Questions were played perfectly - so that each yes or no yields the contestant one bit of information - then twenty bits would suffice to identify 2^{20} candidate - objects. But this power requires that there must be nothing random about these questions and they require vast background knowledge for the careful partitioning of the information - space at every stage." Just as in Twenty Questions so it is in Mastermind and any deductively organized science, participants are advised to avoid random or reckless reasoning. Quite the contrary, there is every reason to rehearse and share with others the arguments that lead one to endorse one hypothesis rather than another prior to soliciting a response. For instance, in the present example of testing H2 above, a child may reason:

If I find that T2 fails to confirm any of the components of the hypothesized patterns, then I have learned that three colors are eliminated altogether! Furthermore I will have learned that NL is made up of either yellow, and/or white pegs with at least one of the two repeated at least once. Of course, there are many other "If . . . then . . ." statements the child may utilize and teachers need not go into each of them. The point is simply that these thought experiments of logic accelerate the accumulation of empirical information in any deductively organized research tradition.

Once a system is bounded with appropriate defining postulates, the first hypothesis selected for testing is something of an educated guess. But this educated guess is a bit more than a stab in the dark.

The Eutochia

The first hypothesis for a test is singled out at the ILCS for special attention. H1 is referred to as a eutochia. Eutochia is a term taken from the ancient Greek and means educated guess. In Mastermind, the eutochia is both informed by and constrained by the postulated set of colors, and the stipulation that NL must contain three immediately sequential relations comprised of no more than, at most, four named color pegs (i.e. four semantic referring terms and their syntactically imposed relations).

The job of the eutochia is to reveal sufficient information upon which to base a strategy for future research (Lesher, 1978). Each *optimally* plausible hypothesis thereafter is more strictly constrained by the evidence (Lakatos, 1970). In Mastermind, only recklessness leads one to test an H2 overlapping in content to H1. With each subsequent test, researchers deduce and develop a "feel" for where the research programme is likely to take them (Faust, 1984; Scheffler, 1981; Miller, 1984; Wimer, 1977, 1979). Subsequent to testing each hypothesis, real scientists (Kuhn, 1962) should have a better "feel" for the potency of a research programme (And in an empirically based science they may even decide "not to play this game" as framed. They may instead conceive a different closed-ended operating theatre to explore.). In any case, as long as the experimental process of hypothesis, test, publication and subsequent testing continues within a bounded paradigm, the uncertainty and amount of the unknown reduces by clearly evident fits and starts.

Record-keeping and Publication

Science is a community process. It advances most efficiently when members are informed of the stepwise success of one or another researcher. Moreover, science plods along securing justification for each claim to know rather than grasping at novel observations and then cavalierly speculating as to their meaning. This heuristical need for justification is illustrated to students nicely in Mastermind. Teachers insist from the test of the eutochia onwards that students specify what it is that they know in response to each test of a hypothesis. Moreover, teachers insist that student/scientists not only declare what they know but explain as well how they know what they claim to know. In short, students are required to create an argument hooking experimental observation together with the defining postulates of the game to infer the conclusion of an argument. Each student/scientist must learn how to ask and respond to the following two questions when completing an experiment: “How do I know?” and “How do I know that I know?”

Answering those questions requires patience and skill. In most cases, the younger the student/scientist the more likely she is scrambling to arrange four more pegs for testing even as she is telling what she presumably knows. “You know we got one color right!”, she may report. But that is too cursory . The teacher baits the child further to ferret out all the information available as a result of T1.The teacher must ask (usually several times over), “Is that all you know?”

The actual amount of information recoverable from T1 is too much for people to commit to working memory. The amount of information is astonishing to students who, each in their own way, usually identify only two or three obvious bits of information when first sizing up what has been learned. The need for publicly recording test results begins to emerge.By publishing everything that has been learned on a blackboard or overhead, everyone can see both the extent and the detail of all that has been discovered. By neatly exhibiting all that has been learned as a result of each test, the inferential trail of scientific inference up to that point becomes more evident. Children see an unfolding of the known from experimental data and how it can be organized to model the phenomenon under investigation. From this, children acquire a feeling for the cognitive procedures of science (Lawson, 2005; Scheffler, 1978). This feeling of understanding is often missing in less hands-on activities, or activities that may be "hands-on" but so distracting because of the gadgetry involved, that little attention is directed to the critical inferential process. People, unlike computers, must have a feeling for the inferential process before concluding that they understand it (Dreyfus and Dreyfus, 1986).

True understanding in any field of study occurs only when people acquire the appropriate feelings of understanding along with appropriate knowledge and skills. Perhaps this is the sort of thing Stephen Norris (1985) has in mind when he talks of scientific sophistication. Norris' sophistication transcends anything accomplished through blind adherence to a set of behavioral objectives specified ahead of time (Doll, 1993; Donovan, 1992). A well-scaffolded and hands-on experience, such as the analogical use of Mastermind, leads students to integrate a sense of science into their more general reflective practices as individuals (Chin, Brown, & Bruce, 2002).

Thought Experiments and the Good Test

In Mastermind, rather than set four pegs in place at random, the student/scientist may conclude that the most vigorous eutochia places four pegs all of one color. After all, if that color is eliminated this strategy decreases the uncertainty inherent in the conceptual space by nearly half. That is to say, from 1296 possible codes, NL is reduced to one of approximately 600 alternatives. On the other hand, if the eutochia is comprised of four separate colors and none of them turns out to be in NL, the reduction achieved is even greater. But how likely is it that NL might have only two colors? Which is the most prudent way of deciding on a eutochia? Is it all just a gamble? What ever one decides, the reduction of the unknown *can* occur as a geometric progression as might be demonstrated to more advanced students. Students of all ages however can readily appreciate the fact that some tests are better than others which is to say, some tests reveal more than others. This difference in test quality becomes ever more evident with each experimental cycle, subsequent to the test of a eutochia. The more effective the experimental cycle the less is left unknown at its completion.

The meaningfulness of test results depends both on the power of the hypothesis tested and the skill of the scientist interpreting the results. The power of a hypothesis is built upon three factors: range, specificity and novelty. By range is meant that a hypothesis boldly refers to as much phenomena as can be precisely measured with the possibility of a single disconfirming test (Popper, 1958). Specificity requires that semantically a hypothesis is free of equivocation and ambiguity. It means further that every observational description that ought to be identified as exemplary of the phenomenon under study is identified and confusion with distracting phenomena avoided. Finally, novelty means that a hypothesis strikes out to new territory and doesn't replicate what is already known. Hypotheses replicating territory already covered by previous research are costly in terms of effort and time expended and in sciences involving massive inputs of capital for each experiment, economically unfeasible. Scientists running cyclotrons or setting up a test in Mastermind can ill – afford waste when setting up the next hypotheses for test. To avoid replication of effort, scientists read all that is already known about a subject and create thought experiments *before* subjecting a new agreed upon hypothesis to testing. Such thought experiments are hypothetical deductions treating what is thought to be known already as premises. By following out the deductions, scientists anticipate what can be learned should one hypothesis be tested as opposed to some other. In the example above, when the children conclude T2, they still have no new information about white pegs. Further, since both red and green alone are used in the second test some of the information they in fact acquire is redundant. Better planning in advance, through consideration of several thought experiments could have helped avoid this redundancy and hence avoid unnecessary testing. Discussion of such empirical facts leads students to two important insights about hypotheses and testing. The hypothesis selected for testing is not the one most likely to be true; but rather it is the one that regardless of nature's response or the information retrieved, it will be the one superior to any competing hypothesis (Hawkins, 1966; Herron, 1969; 1971). As Ronald Giere (1979) explains, a good test is one that extorts from nature an optimal amount of information-within the confines of a given theory (or what Giere calls a definitional system) and the richest testable hypothesis. The more explicit, bold, far ranging and novel a hypothesis, the more information can be retrieved from subsequent testing. The more information retrieved the more science has fertile soil for further research.

Another, but not insignificant reason, scientists do not replicate previous experimental work is that there is usually no reward for doing so. Journals do not typically publish articles that say nothing more than that the contents of some previously published article are correct. Mastermind again vividly illustrates the reality of this phenomenon. If several "laboratories" (Mastermind game boards each with its own attendant scientist or group of students/scientists working on figuring out the same NL) are publishing all they know before going on to the next experiment, they see that announcing they have learned NL contains no blue pegs doesn't get published if some other lab has already published the same information. It is well-documented, novel information that gets published, not guesses and not what is already known.

Occasionally, experimental work is replicated. As noted above in the case of Mark Spector, if there is a need for the work product of some previously published experimental procedure then there is replication, not for the purpose of publication, but for the incidental purpose of securing work product. Too, if there is a suspicion that something is amiss in earlier experimental work, scientists may reluctantly take on a replicative study. The reason scientists are reluctant to take on such replicative studies remember is that there is no reward if the studies simply confirm what is already thought to be known. Once again the game Mastermind offers an illustration of both the reluctance to replicate and the need to replicate at times.

On occasion, in the game of Mastermind, Nature may accidentally respond to the test of a hypothesis inaccurately. This is usually regarded as an annoyance of great consequence to people simply sitting around playing the game for the game's sake. But in the context of a class in science education it becomes a wonderful teachable moment lending itself to the currently favored (Taylor, Dawson, & Fraser ,1995) constructivist learning environment. The discovery of the error usually begins when students/scientists begin fumbling around for a powerful new hypothesis to test and given how much has been learned, there seems little to discover. Nevertheless, results from the immediate previous test don't seem to add up with scientist expectations. Eventually students recognize that they have two contradictory confirmations floating about. Since they are dealing with a deductive system they know right off that something must be wrong with one of nature's responses (in real science this shows up as observational error) since contradictions cannot be derived in a closed deductive system without violating a rule of logic. Given enough time and encouragement students usually figure out where the error occurred. Once the correction is acknowledged on the board or overhead and a line drawn through the retracted experimental results the game should continue. Frustrated students will eventually arrive at NL but then there is much fertile ground for exploring the costs of observational error or reporting in real science. (Nature's mistake is treated, in this case, as an observational or a testing error.)

As mentioned above, the point of science is to move away from error and towards truth. Or again, as alluded to above and what should amount to the same thing, the point of science is to continue doing science. Redundant testing of hypotheses is tantamount to doing science by reading a textbook. As long ago as 1903 Montcrief cautioned teachers of science, “The main point for the teacher to keep constantly in mind is that his student is an investigator, seeking by means of his own efforts to find out what is truth- not a mere imitator or verifier of the results obtained by others…The conclusions reached must be deductions from the evidence observed, not statements memorized from a text or learned from a teacher” (p. 351).

In the real world of science (contrary to positivist dictates about experimental evidence confirming or refuting theory (Hempel, 1965; Popper, 1959), conclusions reached can never be anymore than tentative vindication within some theory (Giere,1979).

Elegance and the Language of Science

As evident in the relatively extensive results of T1 it is evident to student/scientists that it is important to record each and every bit of known fact. Further, as noted immediately above, deductions expressed in economical fashion help scientists identify error even in what seem to be mere observations. Scientists collect and *arrange* enormous amounts of experimental information and that alone takes up the majority of their time typically (Gould, 1986).

Arranging information can be overwhelming in the published results of scientists. In the sample game portrayed above it is easy to see that even the results of T1 cover quite a bit of territory if written out in long hand on the blackboard by the teacher. Before discussing science or playing Mastermind, students at the ILCS are taught the rudiments of symbolic logic. This is generally done through the board game Wff n Proof, which uses Polish notation. Students may also be taught directly the operational symbols or

Russellian notation, omitting the bi-conditional, since it plays no role in the sort of deductions children will be involved in and it is ontologically redundant. In any case, with a few simple logical symbols in hand (herein we use Russellian notation) after writing out publishable results in long hand for purposes of illustration, the teacher suggests to students that she wants to make things simpler and take less time recording information. Students are most appreciative of that idea! The trick is to make them feel like co-architects of the process and its emerging language (Lawson,2005). As most teachers of algebra, chemistry and symbolic logic learn early on, show people a statement like: g –> [(bu) V (y)] and people often give up concluding from the outset that they cannot learn "… that kind of stuff". However, as many researchers have shown, more often than not most people can use formalized languages (Garson, 1984; Lawson, 2005; Suppes, 1962)! To learn to use formalized languages people need a "feel" for the problem such languages are meant to address (Wagner, 1985). For example, the teacher in concert with the students, may begin creating a more elegant and formalized language by simply agreeing on some abbreviations. Hence, the first hole to the left of nature (the codemaker) will be identified as A, the second B, the third C and the fourth D. A short while later while recording further test results, the teacher should suggest that the colors be abbreviated as well and in something like the following manner: (b) for black, (bl) for blue, (g) for green, (r) for red, (y) for yellow and (w) for white. And awhile later, the teacher may offer further economies such as NL for nature's law, Tn and Hn for test and hypothesis identification. Finally, a bit later to complete the development of a rudimentary formalized language uniquely suited to the context, the teacher should add the operational symbols from logic. Thus, in the newly created formalized language with precise semantic elements and syntactic rules from standard logic, statement one in T1 above, becomes:

T1

1. (g), -> - [bu V y]

Logical operators are the most important elements in the scientists' vocabulary. "If-then" makes transparent the proposed arguments and claim(s) of hypothesis are made precisely evident with the syntactic functions of "and," "or," and "not." Operations such as "some" are not necessary to employ at this most rudimentary level of explanation nor is there any need to venture into modal operatives such as "probably," "possibly" and so on. Certainly one could not get far explaining inductively derived sciences to students without reliance on such operatives but such explanations are for another day. The "masterminding approach" of the ILCS while enormously effective in explaining the fundamentals of deductive science and other general features of science such as attention to language, publication, elements of experimentation, motivation for replication, quality of tests and so on, it still has its limits as does any other curricular exercise.

The formal languages of science are nothing mysterious. They are just an easy way of keeping track of information. Moreover, the formal language exhibits no more information than is evident in the unwieldy sentences of standard linguistic expression and makes clear the precise nature of inferences being drawn.

Evaluating Scientific Contributions

As the game Mastermind shows, an ill-conceived scientific study produces a minimum of novel information. After students become comfortable with the masterminding process as a model of deductively organized science they are asked to stand outside the process as it were and analyze what went on during the proceedings. Specifically they are asked to reflect upon the contributions made by individual participants to a research initiative beginning with the eutochia. They are then asked to look to all the labs and teams of scientists associated with each lab and consider which contributed the most to the overall advance of science.

The teacher draws everyone's attention to the blocks of information produced sequentially by each lab in its turn. Each block of information should have recorded to the side the name of the lab that produced the results. Students can easily see which lab's experiment produced the most information. But quantity of information isn't all that matters. In real science Nobel laureates are occasionally recognized not on the basis of the extensiveness of their scientific contribution, but rather for the originality and explanatory power afforded by a small but ingenuous contribution that leads to considerable future research. Such was the case with Watson and Crick's Nobel winning article that introduced an apparently accurate model of DNA. Much research had preceded Watson and Crick's' work. But Watson and Crick imagined the currently accepted structure of a double helix and identified sufficient evidence in support of their conjecture (Watson, 1968). Students similarly should consider test results were most critical to finding a hypothesis that mirrored NL.

In the context of such discussions students should be asked to consider a rationale for awarding a prize such as the Nobel in their own masterminding competitions. Students should begin by considering the lab that contributed most significantly to scientific progress. Was it the lab that produced the most information in the nine elements tabulated in response to each team's test of their euthochia? Was it the lab that made the third round of tests of blockbuster significance? Or should the prize simply go to the first lab which produced H4 and the final match with NL?

It does not take long for children to recognize that the lab that confirmed NL may have contributed the least to the development and ultimate completion of the current game (research programme). Somehow significance of contribution seems more adequately measured in terms of imaginativeness and intellectual rigor than by reference to any strictly quantifiable criteria. But how does one go about measuring imaginativeness and rigor in science or Mastermind?

Imaginativeness is an excessively slippery notion and intellectual rigor does not fare much better. However, it is likely that both can be captured somewhat by other scientists with some similar research experience. As Oliver Wendall Holmes Jr. once famously observed in a somewhat analogous situation, "I can't define pornography but I know it when I see it." Scientific colleagues have some sense of what it means to be both rigorous and imaginative within a certain field of science even if no operational definitions come readily to mind.

Identifying the lab that contributed the most is just a first step. Modern science often represents a team effort, but the Nobel Prize is anachronistically awarded for individual accomplishment. At most, only two or three people from a given lab share in the prize regardless of how many scientists worked on the project or how many names appear on the ground breaking publication. And, keep in mind, as on any team of any size, contributions of participants are not easily parsed into equal amounts. The team leader has the final say on which hypothesis is to be tested. Moreover, the team leader is responsible for seeing to it that all information derived from a test is duly recorded, ”published” and shared widely with other researchers (Davidson, 1985; Goodman, 1978; Yager, 2000).

The team leader plays a mighty important role in a team. Besides all that is mentioned above, team leaders may decide to withhold some information from public scrutiny at one time or another, gambling that the lab's next experimental round will lead to such decisive effect that no one can deny the lab's critical role in the advancement of a science. This element of maneuvering for acclaim and reward is never catalogued as a part of "Scientific Method," it is a driving force in what happens (Rudwick, 1985). The team leader plays a central role in all this, but does that mean the leader alone is responsible for the team's success and should alone be awarded the Nobel? Relatively minor team members may prove to have the most logically adept mind; one may have a single but especially potent insight. Obviously if the prize is shared too broadly it will have little meaning in the motivations of scientists. Just how broadly should an award such as the Nobel be shared among members of the most successful team? Making such decisions is central to the hidden decisions and motivations of practicing scientists. Once again, Mastermind brings to light one more teachable moment.

Serendipity

Children know they sometimes identify NL as a result of a lucky guess. Something like this apparently happened at least to some extent in H3 above. Such chance events happen in real science as well. For example, after months of frustrating and unsuccessful research, Louis Pasteur went on vacation and carelessly left behind in his lab some live anthrax bacteria. When Pasteur returned there was no new anthrax to be found immediately at hand so he used the aging, weakened bacteria he had left behind for his next round of experiments. This chance event led happily to the discovery of a vaccination for anthrax. Cows vaccinated by the dying and dead anthrax survived while those vaccinated by new anthrax when it arrived did not. Moreover, when Pasteur vaccinated cows previously vaccinated with old and dying bacteria they did not die even when fresh bacteria were introduced to their system. But, as Pasteur reminded his detractors from the Sorbonne there was more than just chance afoot here. It is only in the presence of a "prepared mind" that a happy accident can be turned into a fortuitous advance in science. In Pasteur's case the accidental use of weakened bacteria was fortunate but more important was that Pasteur was capable of thinking his way through to the fact that dying anthrax stimulated the natural defenses of cattle sufficiently to ward off infection from more virulent strains later on (Harre', 1984).

In the present example of experiment three above, H3's "red, green, red, black" conjecture was surprisingly revealing. As in real science, chance, experiment and logic occasionally combine in Mastermind to bring about rewarding advances in knowledge but only if prepared minds are at hand. Could there be a more convincing lesson to students about the hands on need for self-discipline in both preparatory learning and rigorous reflection?

Research Teams

Teams of children have a natural tendency to look about and see how competing teams are doing. Science is, in most cases, an unavoidably public affair. Similarly in Mastermind with a room filled of competing labs, students are obsessively aware of the efforts of others. Many want to know what others are doing and some others are annoyed by the efforts of others to know what they are doing, “Do your own work!” is not an uncommon exclamation. Competitive scientists may be driven by personal acclaim but the more everyone works together, the more quickly everyone gets the answers respective communities of scientists are seeking. Getting information from other labs is a good thing but the reticence of a lab to reveal information before publication is understandable. More important is within the teams themselves. A brainy sleuth who fails to submit his thought experiments for refutation by others is failing the team. If the team leader relies on his own thinking alone rather than encouraging all to critique each and every proposed thought experiment before a hypothesis is agreed upon for actual test, he is also failing the team. A research team just like any other team wins as a team and fails as a team regardless of the presence of genuine stand out stars or leaders. Considerations of a team’s winning or failing brings forward at this point, yet three more insights into science: namely, what counts as an answer, what counts as a mistake and why scientists share the information they share.

Refutation and Confirmation

An answer in science does not look like the sort of thing a teacher may require on a routine test or the sort of response the ETS folks may demand on the SAT. Whether explicit or not, each scientific explanation and description (i.e. the answers to scientists' questions) is proffered within the context of a theory of error (Achinstein, 2003). Most practicing scientists are well aware that they trade in "near truth and not the Truth of laymen and metaphysicians (Scheffler, 1982).

Answers to a scientific hypothesis come in two forms. First, they may confirm (by failing to refute) a given hypothesis. Second, they may refute a hypothesis outright. Note the importance of this extremely important message for students! A refutation is not a mistake. Refutations tell scientists there is something they surely don't know - given the premises and background knowledge contained in a given research programme. Far from being a mistake, a refutation often provides more certain knowledge than any confirmation of a versimilitudeiness sort ever could. Both are legitimate answers in the practice of science. Mistakes are something altogether different. In science mistakes refer to any event that artificially limits the scientist's attempt to move further away from error.

Furthermore, in science, mistake and error are two separate notions, a distinction too few students understand. Falling subject to a carelessly construed line of reasoning is a mistake in science as elsewhere in human intellectualizing. Again, mistakes distract from the effort to move further away from error. Similarly, failing to perceive the presence of something that should be evident is again a mistake both in science and elsewhere. By contrast, error is something that is inherent in the practice of scientific observation and is due to limitations of the human transducer function or, limitations or malfunction of the technological apparatus employed in a study. Mistakes have no place in science. Their presence is a grievous disappointment. Errors on the other hand are part and parcel of scientific practice and must be accommodated as Kyburg explains within an accompanying theory of error (1984). For example, in two out of three of Eddington's photographic plates there was no confirmation of relativity theory, but Eddington concluded each disconfirming photograph was within an acceptable range of error. In saying this Eddington was acknowledging the perturbations forever destined to contaminate human/observed world interactions. Eddington's observations were not mistaken, but they were outside the exact predictions of Einstein's theory. Answers in science are never error-free. Answers become increasingly acceptable by proving their reliability prompting subsequent revealing research.

Once again Mastermind proves instructive. As noted above there are times when nature may respond to a test inaccurately. The mistake is usually caught somewhere down the line when student/scientists recognize unequivocally that there is a contradiction between one or more published results of tests.

Deductions are truth-preserving techniques of inference-making. They let scientists extend their understanding from well-entrenched commitments of which they are most confident to hypotheses of considerable plausibility (Goodman, 1978; Holloway, 2000). Any conclusions deduced are true--*given* the truth of the premises. Since contradictory conclusions cannot both be true, the existence of some previous observational mistake becomes evident. The contradiction limited investigator's ability to move away from error. Similarly, when considering each of two hypotheses for test, the thought experiments by students generally give only plausible credence to the power of one hypothesis as compared to another when finally subjected to test. There usually is no way to guarantee an optimal revelation of information through review of the thought experiments alone. The variance acknowledged is compatible with having a relevant theory of error in this case.

Conclusion

As noted above, the ILCS has been particularly successful in training teachers in the use of Mastermind for introducing students to some of the subtler aspects of science generally and, more specifically, the deductively organized sciences. The inexpensive masterminding approach to science described above shows students the conceptual architecture of science extends far beyond merely listing facts or watching nature pass by displaying her wares. The approach described herein shows that science is an active engagement with nature aimed at putting her secrets under human control and subject to human fallibility. Few programs in science education reducible to thirty hours of hands on instruction can produce much more positive effect in students.

References

AbuSharbain, E. (2002). Enhancing inservice teacher's constructivist epistemology through the development and redesign of inquiry-based investigations together with their students. *Electronic Journal of Science Education, 7* (1).

Achinstein, P. (2003). *The nature of evidence.* Cambridge: Cambridge University Press.

Achinstein, P. (1985). The method of hypothesis testing: What is it supposed to do and can it do it?" In P. Achinstein & O. Hannaway (Eds.), *Observation, experiment and hypothesis in modern physical science* (pp. 127-146). Cambridge: MIT Press.

Achinstein, P. (1983). *The nature of explanation.* Oxford: Oxford University Press.

Ackermann, R. (1985). *Data, instruments and theory.* Princeton: Princeton University Press.

Adams, P.E., & Krockover, G.H. (1997). Concerns and perceptions of beginning secondary science and mathematics teachers. *Science Education, 81* (1), 29-50.

American Association for the Advancement of Science (1993). Project 2061: Benchmarks for science literacy. New York: Oxford University Press.

Aris, R., Davis, H. T. , & Stuewer, R. H. (Eds.). (1983). *Springs of scientific creativity*. Minneapolis: University of Minnesota.

Ayer, A.J. (1968). *The origins of pragmatism: C.S. Pierce and William James*. San Francisco: Freeman Cooper.

Blackburn, S. (2005) *Truth.* New York: Oxford University Press.

Bridgman, P. W. (1927). *The logic of modern physics*. New York: Macmillan.

Broad, W., & Wade, N. (1984). *Betrayers of truth.* New York: Touchstone Books.

Chin, C., Brown, D.E., & Bruce, B.C. (2002). Student-generated questions: A meaningful aspect of learning in science. *International Journal of Science Education, 24* (5), 521-549.

Cohen, I. B. (1985). *Revolution in science*. Cambridge: Harvard University.

Davidson, D.(2005) *Truth and predication.* Cambridge, Mass.: Harvard.

Davidson, D. (2001). On the very idea of a conceptual scheme. In D. Davidson (Ed.), Truth and interpretation (p. 183). Oxford: Oxford University Press.

Davidson, D. (2001). A coherence theory of truth and knowledge. In D. Davidson (Ed.), *Subjective,* i*ntersubjective, objective*. New York: Oxford University Press.

Davidson, D. (2005). *Truth and predication*. Cambridge: Harvard University Press.

De Boer, G. (1991). *A history of ideas in science education.* New York: Teachers College Press.

Doll, W.E.(1993). Curriculum possibilities in a “post” future. *Journal of Curriculum and Supervision, 8* (4), 277-92.

Donovan, J. F. (1992). The Hunter model: A four year longitudinal study of staff development effects. *Journal of Research and Development in Education, 25*, (3), 165-72.

Dreyfus, S.E., & Dreyfus, H.L. (1986). Mind over machine. New York: Free Press

Dummett, M., (2004). *Truth and the past*. New York: Columbia University Press.

Ericcsson, K.H., & Simon, H. (1984). *Protocol analysis: Verbal reports as data.* Cambridge: MIT Press.

Bernard, d’Espagnat (1983). *In search of reality.* New York: Springer-Verlag.

Faust, D. (1984). *The limits of scientific reasoning.* Minneapolis: University of Minnesota.

Frank, P. (1947). *Einstein, his life and time.* New York: Alfred A. Knopf.

Frykholm, J. (2005). Connecting science and mathematics instruction: Pedagogical context knowledge for teachers. *School Science and Mathematics, 105* (3), 127-41.

Gallagher, J.J. (2000). Teaching for understanding and application of science knowledge. *School Science and Mathematics, 100* (6), 310-18.

Garson, J. (1980). Teaching logic with Emil. *Teaching Philosophy,* 3, 453-478.

Goodman, N. (1978). *Ways of worldmaking*. Indianapolis: Hackett.

Gould, S. J. (1986). A triumph of historical excavation. *New York Review of Book (*XXXIII) 3, 9-15.

Giere, R. (1979). *Understanding scientific reasoning.* New York: Holt Rienhart and Winston.

Hammrich, P.L. (1997, May/June). Teaching for excellence in K-8 science education: Using Project 2061 benchmarks for more effective science instruction*.* *Journal of Teacher Education, 48* (3), 222-232.

Haney, J., Czerniak, C. M., Lumpe, A.T. (2003, December). Constructivist beliefs about the science classroom learning environment: Perspectives from teachers, administrators, parents, community members, and students. *School Science and Mathematics, 103* (8), 366-77.

Harre', R. (1984). *Great scientific experiments.* New York: Phaidon.

Hawkins, D. (1966). Education and the spirit of science-critique of a statement. *Science Teacher,* 6, 105-107.

Hempel, C. (1965). *Aspects of scientific explanation.* New York: Free Press.

Herron, M. D. (1969). Nature of science: Panacea or a pandora's box? *Journal of Research,* 33, 18-20.

Herron, M. D. (1971). The nature of scientific inquiry. *School Review*, 79, 170-212.

Holloway, J.H. (2000). How does the brain learn science? *Educational Leadership, 58* (3), 85-6.

Horwich, P. (1990) *Truth.* Oxford: Blackwell.

Hug, B., Krajcik, J., & Marx, R. W. (2005). Using innovative learning technologies to promote learning and engagement in an urban science classroom. *Urban Education*, *40* (4), 446-472.

Koertge, N. (1978). Towards a new theory of scientific inquiry. In G. Radnitzky, & G. Anderson (Eds.), *Progress and rationality in science*. Boston: D. Riedel.

Kuhn, T. (1962). *The structure of scientific revolution.* Chicago: University of Chicago.

Kyburg, H. J. (1985). *Theory and measurement.* Cambridge: Cambridge University Press.

Lakatos, I. (1970). Falsification and the methodology of scientific research programmes. In I. Lakatos & A. Musgrave, (Ed.), *Criticism and the growth of knowledge*. Cambridge: Cambridge University Press.

Lawson, A.E. (2005). What is the role of induction and deduction in reasoning and scientific inquiry? *Journal of Research in Science Teaching, 42* (6), 716-40.

Lesher, J. H. (1978). On the role of guesswork in science. *Studies in History and Philosophy of Science,* 9, 1.

Lynch, M.P. (2005). *True to life*. Cambridge, Mass.:MIT Press.

Miller, A.I. (1984). *Imagery in scientific thought.* Boston: Birkhauser.

Montcrief, W. F. (1903). A plea for experimental work by the student in teaching a first course in physics. *School Science and Mathematics, 3* (6), 349-354.

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

Nisbet, J. (1993). The thinking curriculum. *Educational Psychology, 13*, 281-290.

Norris, S.P. (1985). The philosophical basis of observation in science and science Education. *Journal of Research in Science Teaching*, 22 (9), 817-833. (An earlier version of this paper was presented at the First Annual conference of the Institute for Logic and Cognitive Studies, June 12, 1984).

Oliver, J.S., & Nichols, B. K. (2001). Intellectual independence as a persistent theme in the literature of science education: 1900-1950. *School Science and Mathematics, 101* (1), 49-56.

Pais, A. (1982). *Subtle is the Lord…The science and the life of Albert Einstein*. Oxford: Oxford University Press.

Papineau, D. (1979). *Theory and meaning*. Oxford: Oxford University.

Pierce, C.S. (1905). *Collected Papers*. (Eds.), C. Hartshorne, P. Weiss, & A.W. Burks. Cambridge: Harvard University.

Popper, K. (1959). *The logic of scientific discovery*. New York: Harper Row.

Popper, K. (1972). *Objective knowledge: An evaluationary approach.* Oxford: Oxford University.

Putnam, H. (1975). What is mathematical truth?, in *Philosophical Papers Vol.I.* Cambridge: Cambridge University Press.

Rescher, N. (1978). *Pierce's philosopohy of science.* Notre Dame, Indiana: University of Notre Dame.

Rudwich, M.(1986). *The great devonian controversy: The shaping of scientific knowledge among gentlemanly specialists*. Chicago: University of Chicago.

Salmon, W. (1984). *Scientific explanation and the causal structure of the world.* Princeton: Princeton University.

Salmon, W. (1967). *The foundations of scientific inference.* Pittsburgh: University of Pittsburgh.

Scheffler, I. (1978). In praise of cognitive emotions. *Teachers College Record, 79* (2)171-186. Columbia University, New York, NY.

Scheffler, I. (1981). *The anatomy of inquiry.* Indianapolis: Hackett.

Scheffler, I. (1982). *Science and subjectivity.* Indianapolis: Hackett.

Schlick, M. (1925). *General theory of knowledge*. Berlin: Julius Springer.

Stenhouse, D. (1985). *Active philosophy in education and science*. London: George Allen and Unwin Ltd.

Suppe, F. (1977). *The structure of scientific theories*. Urbana: University of Illinois.

Suppes, P., & Hill, S. (1962). Mathematical logic for the schools. *The Arithmetic Teacher,* 9, 369-399.

Tarski, A. (1965). *Introduction to logic and to the methodology of the deductive sciences.* Oxford: Oxford University Press.

Taylor, P., Dawson, V., & Fraser, B. (1995, April). *Classroom learning environments under transformation: A constructivist perspective.* Paper presented at the meeting of the American Educational Research Association, San Francisco, CA.

Toulmin, S. (1953). *Philosophy of science.* London: Hutchinson.

UNESCO (1980). UNESCO handbook for science teachers. Paris, UNESCO/London: Heinemann.

Van Fraasen, B. (1980). *The scientific image*. Cambridge: Cambridge University Press.

Wagner, P. (1983). The nature of paradigmatic shifts and the goals of science education. *Science Education, 67* (5), 910-924.

Wagner, P. (1985, October). The language of science. Paper presented at the second annual conference of the Institute for Logic and Cognitive Studies, Houston, Texas.

Wagner, P. (2006). Teaching and understanding science. Manuscript in Preparation.

Williams, B. (2002) *Truth and truthfulness*. Cambridge: Cambridge University Press.

Watkins, J. (1984). *Science and skepticism.* Princeton: Princeton University Press.

Watson, J. (1968). *The double helix*. New York: Atheneum Press.

Weimer, W. B. (1977). Scientific inquiry, assessment, and logic. *Therapy and Research, 1*, 235-258.

Weimer, W. B. (1979). *Notes on the methodology of scientific research.* Hillsdale, New Jersey: Earlbaum.

Zahar, E. (1978). Crucial experiments: A case study. In G. Rudnitzky & G. Anderson (Eds.), Progress and rationality in science. Boston: D. Reidel Publishing.

Venvitt, G., Adey, P., & Larkin, S. (2003). Fostering thinking through science in the early years of schooling. *International Journal of Science Education, 25* (11), 1313-1331.

Yager, R.E. (2000). A vision for what science education should be like for the first 25 years of a new millennium. *School Science and Mathematics, 100* (6), 327-41.

About the authors…

Paul A. Wagner is a former Executive Secretary of the Philosophy of Education Society, former Vice President of the Association of Philosophers in Education, president of the Texas Educational Foundations Society and member of the Philosophy of Science Association. He has also served as an adjunct professor in the School of Natural and Applied Sciences wherein he has taught a course titled Development of the Sciences.

Lillian Benavente-McEnery is an assistant professor in the literacy, language, and library science department and Director of the University Reading Clinic in the Center for Educational Programs at UH-Clear Lake.