Chapter 3

An Integrative Analysis of a Problem-Solving Classroom

The focus of this chapter is the analysis, using my model for an 'integrated scientific method' (ISM), of an innovative high school genetics course. The immediate goal of this analysis is to develop a more sophisticated understanding of one science course. A second goal, one with implications that are more generalized and long-term, is to test the analytical utility of ISM and to search for ways to improve ISM as a multi-purpose tool that can be used not just for describing scientific methods, but also for analyzing and designing instruction. Using ISM for classroom analysis will broaden the scope of ISM-based description so it includes both science (in Chapter 2) and instruction (in Chapter 3). Hopefully, this analysis will stimulate creative thinking about the connections between science, ISM, and education, and will serve as a bridge to the development of further educational applications.

The next part of the chapter describes the reasons for selecting the Monona Grove (MG) genetics course for analysis, and the philosophical foundations for the course. Following this is a description of the methods used in the analysis of the MG course. The first phase of the ISM-based analysis examines the 'science experiences' that occur during each of eleven instructional activities in the MG course. The analysis of the MG course concludes by examining the functional relationships between instructional activities, and the resulting 'structure of instruction' in the course. Finally, there are suggestions for how to improve the MG course, and a critique of my analysis that asks whether the analysis achieved its objectives.

3.11: Selection of a Course for Analysis

Why has this particular course been selected for analysis? First, the course already has been studied by a number of researchers, so in my analysis I can use the data they have gathered and the reports they have written, and I can build on the insights they have gained. Second, and more important, this course — to a greater degree than in most courses — gives students an opportunity to experience a wide range of the 'methods of science'.

Why might a wider range of experience be educationally useful? To encourage a wider scope for education, Perkins & Simmons (1988) describe a model of learning with four frames of knowledge (italicized below), and recommend that instruction should include all four frames. But most science classrooms focus on only two of the frames — *content* (theory-**learning**) and, to a lesser extent, conventional *problem solving* (theory-**using**) — so most students learn little about knowledge and skills in the *epistemic* and *inquiry* frames, about modes of thinking that involve theory-**evaluating** and theory-**revising**. The inquiry frame is the most frequently neglected, partly because it is "the most ambitious and perhaps hardest to cultivate through education. (Perkins & Simmons, 1988, p. 313)" But scientific inquiry is the main focus of Sue Johnson's genetics course. Giving students an opportunity for an exciting 'science in action' experience is one of the main course objectives, as described by the course developers:

> A good knowledge of science involves experiencing first-hand the production and application of scientific knowledge. ... [In the MG classroom] students work in research groups to tackle problems, build models to explain phenomena, and defend and critique those models. ... The methods they use are those of the research scientist. (Johnson & Stewart, 1990; pp. 298, 306)

Based on a preliminary review of the literature describing it, this course seemed to give students significant opportunities for in-depth experience in many of the methods of science. Based on this expectation, I selected the MG course — which encompasses a wide range of scientific methods, with relatively few "blank spots" where essential activities of science are missing — because I thought it would provide a good context for exploring and testing the analytical utility of ISM. Compared with conventional instruction, there would be a wider-than-usual range of instructional activities to be creatively analyzed within the framework of ISM, so there would be more possibilities for the stimulation of productive ideas about ISM and its potential applications for education.

3.12: A Classroom Context for Problem Solving

This section describes the philosophical and pedagogical foundations that have guided the design of the Monona Grove genetics course. Its two subsections examine the connections between problem solving and learning, and the implementation of educational philosophies in the MG genetics course. The papers discussed in this section were written by the designers of the MG course: James Stewart, Susan Johnson, and Robert Hafner.

In science and philosophy the word 'model' is used to represent a variety of related yet subtly different meanings. In Chapter 3 the intended meaning can be understood by considering the context. Often, to conform to a common use of this term in previous literature for the MG course, the terms 'model' and 'theory' will have approximately the same meaning.

A. Effect-to-Cause Problems

Stewart (1988) explains why different types of problems may differ in their potential to promote different types of learning outcomes. This paper describes: A) different *learning outcomes,* B) different *problem types,* and C) *interactions* between problem types and learning outcomes.

Learning Outcomes. By solving problems, students can gain a better understanding of five types of knowledge: the conceptual structure of a discipline, general problem-solving heuristics, discipline-specific instantiations of general heuristics, discipline-specific problem-solving algorithms, and the nature of science as an intellectual activity.

Problem Types. Stewart suggests that instead of categorizing problems by content (for example, by the kind of genetics inheritance pattern being studied), a typology based on "type of thinking" might be more useful in promoting the five learning outcomes. In genetics there are two main types of problems, requiring two types of thinking. In cause-to-effect problems, a solver reasons from a cause (such as a known inheritance pattern) to predicted effects (about the visible characteristics of offspring). An effect-to-cause problem requires a solver to reason from observed effects (data about the offspring) to a cause (an inheritance pattern that can explain the data). Most problems in introductory textbooks (for high school or college) require *cause-to-effect* reasoning, and can be solved using content-specific algorithms. Problems that require *effect-to-cause* reasoning involve the interpretation of effects (observations) that are provided for students (as in a textbook problem) or that students must generate for themselves (in a lab or with a computer simulation).

Interactions between Problem Types and Learning Outcomes. The author explains some ways that effect-to-cause problems promote learning — for example, by providing motivation and opportunity for students to develop a well-organized base of conceptual knowledge, including the use of general and genetics-specific algorithms and heuristics. In addition, effect-to-cause problems more closely simulate typical scientific research, so in doing these problems the students are more likely to learn valuable lessons about science.

B. The Classroom

The essential idea from above is that effect-to-cause problems can help students learn the skills of science and the nature of science. In a related paper, Stewart & Hafner (1991) propose that problem-solving research should be extended to activities related not just to a 'context of justification', but also to a 'context of discovery'; school science should include not just *what* is known by science, but also *how* this knowledge is generated. These two ideas — about effect-to-cause problems, and the generation of scientific knowledge — are part of a broader instructional philosophy that has guided the design and operation of the Monona Grove genetics course, as described in papers by Johnson & Stewart (1990) and by Stewart, Hafner, Johnson, & Finkel (1992). This course will be discussed in terms of interactions between three elements: a computer simulation, a model of science, and classroom instruction.

Genetics Construction Kit (GCK) is a computer simulation of a laboratory where scientists do genetics research. The development of GCK was motivated and guided by an educational philosophy:

> I sought a way to expose more of my students to the pleasures, frustrations, and the logic of long-term, extensive research problems where they could react to information collected at any stage in a series of experiments and react accordingly. ... [This] computer program creates an experimental universe in which most of the professional tools are available to perform open ended experiments. (Jungck & Calley, 1985, p. 12)

GCK is compatible with a '3Ps' view of science as Posing problems, Probing problems, and

Persuasion of peers (Peterson & Jungck, 1988). In fact, there is a personal connection between GCK and 3Ps: John Jungck, co-developer of GCK, is also co-author of the original paper describing the 3Ps view of science and science education. Applied to education, a 3Ps philosophy says that students should have opportunities to experience the process of science. And the GCK program provides this type of opportunity. GCK and the 3Ps are essential components of the Monona Grove genetics course:

> The real significance of the computer was that it allowed the classroom to become a place where students learned genetics by engaging in many of the important activities of geneticists. ... The computer use was guided by a view of science and science teaching as problem posing, problem probing, and persuasion of peers. ... (Stewart, Hafner, Johnson, & Finkel, 1992, p. 334)

The most important feature of the MG course is that it gives students an opportunity to engage in many activities of research science. GCK performs a valuable function by letting students do effect-to-cause problem solving that involves not just using models, but using-and-revising models. But providing opportunities for scientific experience does not guarantee that students will learn from their experience. Although the main theme of Stewart (1988) is that solving problems can promote learning, the importance of instructional support is recognized:

> Problems only create a situation where there is the potential for realizing learning outcomes. Whether or not they are realized will be a function of the instructional environment, particularly teacher expectations, in which the problems are set. (Stewart, 1988, p. 238)

One important part of this environment is getting students to "think about what they are doing" in metacognitive reflection. Another aspect, one that is especially appealing to students, is the contagious enthusiasm of the teacher and a general orientation toward having fun. In this course, when things get rolling the ambience of the classroom — as an exciting "place for doing science," as a setting for the intellectual adventure of discovery — produces a rich context for learning:

> The classroom took on the atmosphere of a professional scientific community; the language of the classroom was genetics. New teacher and student relationships emerged. The teacher became the senior researcher, a collaborator, and the students the junior researchers. The teacher's authority came not from the knowledge of solutions but from her expertise as a strategist in solving open-ended problems. (Stewart, Hafner, Johnson, & Finkel, 1992, p. 334)

There are also significant intellectual challenges. With the effect-to-cause problems posed by GCK in this course, students must do a different type of cognitive work (not demanded by typical problems) in the areas of posing, probing and persuasion, because the problems generated by GCK are incompletely structured, thus forcing students to do some of the ‘structuring work’ that is common in scientific research, but is usually done for students by the textbook or teacher.

The overall aim of the genetics course is to help students achieve all of the learning outcomes described by Stewart (1988):

> By situating the conceptual knowledge of genetics in the context of its use, we felt that the students would develop a highly structured and functional understanding of that knowledge, general and genetics-specific problem-solving strategies, and insight into what it is like to ‘do’ genetics. (Johnson & Stewart, 1990, p. 297)

The main theme of "Using Philosophy of Science in Curriculum Development..." (Johnson & Stewart, 1990) is that the MG genetics course has been consciously situated in the context of a philosophy (including 3Ps and more) that the course designers feel is especially conducive to learning both process and content. This philosophy differs from that in conventional content-oriented science education:

> [With conventional objectives] there doesn't seem to be time to teach all the ‘important’ content. But, because of that, we may have missed some of the best lessons of science — those having to do with its very nature. Has it been too easy to omit these lessons from the curriculum because the nature of science is seen as a separate topic rather than a common thread? [By contrast, a 3Ps perspective of science education] does not see a dichotomy between the learning of scientific content and learning about the philosophical nature of science; rather, it sees them as mutually supportive. (Johnson & Stewart, 1990, p. 298)

3.2: Methods for the Analysis

Sections 3.21-3.28 describe how the analysis was done, beginning with the overall strategy, continuing with a characterization of the classroom, and concluding with a description of the three stages of analysis and the data sources that were used.

3.21: Activities and Experiences in a Functional Analysis

ISM was constructed in the form of a descriptive framework that is intended to be useful for describing and communicating ideas about science. An important part of this framework is its visual representation, the *ISM diagram.* For Objective 2, an ISM-based analysis of the Monona Grove classroom, this diagram will be supplemented by a new type of diagram, intended to be useful for the analysis of instruction, that is a grid with science experiences (based on the ISM model) on one axis and student activities on the other axis. A grid can be used to keep a record of the experiences that are available during each activity. Here is a simple version of an *activity-and-experience grid:*

Table 1 (this is in the "tables.pdf" file)

This grid clearly shows *multiple experiences* (for example, scanning vertically down the column for Activity #2 shows that this activity provides Experiences B and C) and *repeated experiences* (scanning horizontally across the C-row shows that experience with C occurs during Activities 2, 3 and 5). A distinction between experiences that are 'multiple' or 'related' can be made in terms of time and coherence. If many experiences occur during a relatively short interval of time, and if they are coherently related to a single unifying theme (as, for example, when doing a one-day laboratory), they will be treated as 'multiple experiences' that occur during a single activity. But if one coherent group of experiences is separated from another group of experiences by a longer interval of time, the two groups will be treated as two distinct activities (such as #1 and #2) and there is a possibility for 'repeated experiences'. Of course, these general guidelines do not always define a sharp boundary between different activities. And sometimes it may be useful to think in terms of smaller sub-activities within an activity, and coherently related activities that combine to form a larger mega-activity.

By increasing the size of the cells in a grid, during analysis there will be room for comments that are more informative than the simple "yes" or no ("—") shown above. Whatever information is considered useful can be included in the body of the grid. It may seem useful to distinguish between major and minor activities on the basis of how much time students invest in an activity, or whether it is crucial for determining success in problem solving, or whether it is a "rare opportunity" that is present in the instruction being studied but is not usually available in other instruction. Or an experience may be repeated several times during one activity. Or a certain type of problem-solving experience may be especially difficult for students, so during the course there may be an attempt to gradually build up student skills in this area, and to gradually increase the difficulty level of problems that require this skill. Any of these characteristics, and others, can be described in a cell.

The visually meaningful organization of information in a grid — such as seeing multiple experiences in a vertical column, and repeated experiences in a horizontal row — can promote an improved understanding of the pedagogically functional relationships between experiences, between activities, and between activities and experiences. These functional relationships include multiple and repeated experiences, the sequencing and overlapping of activities to form 'spirals' within the course, and more.

A grid also shows gaps that can orient the design of new activities. For example, after Activities 1-3, students still have no experience doing A; when this gap becomes known, a teacher may decide to add Activities 4 and 5. In many cases a new activity is a modified version of an old activity; for example, Activity 2 might be an expanded version of Activity 1, revised so it also includes Experience C. During Activities 1-4 students can experience A-D, but occasionally it may be beneficial to use a single activity, such as #5, that by itself provides students with the opportunity for a more complete range of experience. This type of "check list" approach to using an ISM-based grid should be considered a first level of analysis, the most basic approach. This first level is based on the assumption that an education in the processes of science should include, as a minimum requirement, an opportunity to experience these processes. In determining whether opportunities for experience exist, using an activity-and-experience (A-and-E) grid as a simple check-list can be helpful. But, as described above, a grid can also be used to facilitate a more sophisticated analysis, with the goal of gaining a deeper understanding of the functional relationships between experiences, between activities, and between activities and experiences.

In my analysis of the MG classroom, the goal was to understand the structure of instruction, and its relationship to student experiences. But I have not attempted to describe the learning outcomes for students, because a determination of learning would require assessment procedures beyond the scope of my analysis. Similarly, my analysis was intended to examine the structure of instruction, not to evaluate the effectiveness of instruction.

3.22: An Overview of the Analysis

Before describing the analysis it will be useful to make a distinction between student activities and student actions. I am defining a *student activity* as "what students are asked to do," and *student actions* as what students really do in response to what they are asked to do. For example, if students are asked to construct an explanation for the black box, this request (to solve a problem) defines the student activity. When students pursue a solution to this problem, what they actually do are 'student actions', which are thus defined analogous to the problem-solving 'scientist actions' in ISM.

The immediate goal of analysis is to gain a deeper understanding of the *structure of instruction* in the MG genetics course. The overall process of analysis can be summarized in a four-step process:

1) Before the MG analysis began, I studied the actions of scientists in an effort to gain a deeper understanding of these activities and their integrated relationships, and used this understanding to construct the ISM framework.

2a) My pre-analysis understanding of scientist-actions, described in the ISM framework, was used to define "science experiences" at the left border of an activity-and-experience grid.

2b) My pre-analysis understanding of student activities was used to define the "student activities" at the top border of an A-and-E grid.

3) During the MG analysis the ISM-based grid, constructed in Steps 2a and 2b, was used to analyze student actions (what students actually do) in order to gain a deeper understanding of how these are related to student activities (what students are asked to do). For example, the analysis examined the ways that student actions are related to the scientist actions summarized in the "science experiences" border of the A-and-E grid, and asked how opportunities for student actions (and thus the science experiences) are produced by the 'structure of instruction' (including the planned activities, and the teacher's improvised 'coaching' during these activities) in the classroom.

The result of Step 1 is the ISM framework (in Sections 2.01-2.07). Steps 2a and 2b, which define the analytical categories in the A-and-E grid used for analysis, are described in Sections 3.26 and 3.23-3.25, respectively. The methods used in Step 3, the analysis, are discussed in 3.27-3.28.

3.23: Major Instructional Activities

This section describes the major instructional activities that appear on the top border of the A-and-E grid used for the analysis. In the following description by the teacher (Johnson, 1996), instructional activities have been divided (by myself, with group labels such as "1. Black Box Model Revising" added by me) into five groups:

1. **Black Box Model Revising**: The students, organized into self-selected research groups consisting of three members, begin with an activity in which they are presented with a detergent container that dispenses a preset amount of liquid. Working in their groups, they are to build an explanatory model in response to observations they make and to draw a picture of that model on a large piece of construction paper. The second day a scientific classroom conference is held and each group presents their model(s) to the other groups so that it might be critiqued. The groups spend the third day revising their models in response to the critiquing process and then present their revised models, which are taped to the wall next to their original models. This activity is done to have the students experience the process of model building and the subsequent process of model revising. Through this activity they use strategies to produce and assess their models.

2. **Genetics Phenomena**: This initial activity is then followed by a series of others in which the students are introduced to the types of phenomena, such as similarity, diversity, and continuity, for which geneticists attempt to build models. These activities include a lab, *Food for Thought*: *The Cookie Analogy,* in which the students are assigned a certain type of cookie to bake and bring to class. The analogy of recipes for cookies and "recipes" for organisms is used to stimulate discussion on genetic phenomena. In the second activity the students determine their phenotype for a limited number of traits and variations in order to observe similarity and diversity in *Homo sapiens*. In the third activity, the students study human pedigrees and compile a list of phenomena for which geneticists would need to build explanatory models. These phenomena might include offspring that express a variation of a trait not observed in either parent or a particular variation that is found more commonly in males than females.

3. **Initial Models**: The students then read Mendel's paper in which the model of simple dominance is presented in an attempt to begin to explain those phenomena. A graduate student dressed as Mendel visits the class and provides the teams with pea flowers and three generations of peas. With Mendel they count and classify peas and reconstruct the process through which he produced the model of simple dominance. Working with Mendel they are introduced to a structured way of thinking about that model, including the objects (i.e. genes and alleles), states in which those objects can exist (i.e. homozygous or heterozygous), and finally processes through which the states are changed (i.e. segregation and independent assortment). ...

   The student research groups become even more familiar with the simple dominance model as they work with populations of "fruit flies" provided by the computer simulation GCK. ... As they [use GCK] they are asked to think about them in terms of the objects, states and processes found within Mendel's simple dominance model.

   Next the students review the mitotic model of nuclear replication and division. This model has been constructed by scientists to explain asexual reproduction, which involves only one parent and produces offspring genetically identical to that parent. The student research teams then revise the mitotic model in order to produce one that can be used to explain sexual reproduction. Class discussion of the models produced is used to construct a single working model of meiosis.

The groups become more familiar with the model by using it to explain subsequent crosses made using GCK. At this point each group is tested as a unit to demonstrate their knowledge of the simple dominance and meiotic models. The test includes one and two trait problems generated by GCK. The groups must decide what information can be determined from each of the vials, including whether a particular allele is dominant or recessive, and the possible genotypes of parents and offspring in that vial. They also use meiotic models to give evidence for their answers at the chromosomal level.

4. **Genetics Model Revising**: Following the test the students begin three rounds of model revising. In each of these rounds they are presented with a population of "fruit flies", generated by GCK, which when crossed produce data anomalous to the simple dominance model. The students attempt to detect the anomaly and then work to revise the simple dominance model to explain that anomaly. In this effect-to-cause problem solving, similar to that done by classical geneticists, the problems are open ended with no "correct" solution as multiple models can be proposed and assessed.

   After two to three days of each round of model revision, the groups who have produced working models, present those models at a classroom conference. The presenting group is given a population of flies (using GCK) that exhibits the type of inheritance pattern with which they have been working. Using a projection plate, so that the other teams can observe the process, they make two to four crosses to apply their model to that particular population. At this point the other research groups present at the conference can suggest a particular cross. The presenting group has to predict what the results of that cross will be. It is one of the most powerful moments in the class when a group teaches their peers a model that they have constructed and use it to predict the results of any cross. Not only are their classmates impressed with what the presenters have accomplished, but they can see one of the important uses of a model, to make predictions about, as yet, unobserved phenomena. ...

5. **Manuscript Preparation**: The groups each choose one model that they have constructed and prepare a draft manuscript describing both the model and the processes involved in its construction. The manuscripts are then submitted to the editor (in this case, the classroom teacher). The teams revise their manuscripts in response to the editorial critiques and then the final draft is submitted for publication in *The Proceedings of Monona Grove Science.* (Johnson, 1996, pp. 33-35, 37, 39)"

The five activity groups, described above, contain the following major activities:

1. **Black Box Model Revising**: Constructing Models (for the Black Box), Conference (to discuss models), and Revising Models.
2. **Genetics Phenomena**: Cookie Analogy, Human Variations, and Human Pedigrees.
3. **Initial Models**: Read Mendel's Paper, Visit by Mendel and Construction of a Mendelian Model, GCK Practice with Mendelian Model, Review Mitotic Model and revise it to Construct Meiotic Model, and GCK Practice with Mendel/Meiosis Model, Exam covering Mendel/Meiosis.
4. **Genetics Model Revising**: Round 1 of Model Revising, and Conference (to discuss Round 1 models); Round 2 of Model Revising, and Conference (to discuss Round 2 models); Round 3 of Model Revising, and Conference (to discuss Round 3 models).
5. **Manuscript Preparation**: Writing and Revising a Manuscript.

My reasons for selecting these categories, and placing activities within them, is mainly chronological, but there are often logical dividing lines between activities. For the most part, the activities seem to divide easily into categories, with the exception of "GCK Practice" which could fit logically (and chronologically) into Group #3 or #4, or it could be placed in a separate group by itself.

The teacher estimates the following approximate times for each activity:

1. **Black Box Model Revising**: 3 days, with roughly 1 day for each phase.
2. **Genetics Phenomena**: usually about 3 days.
3. **Initial Models**: 7 days total — about 2 days for each model, with 3 days of GCK work between the models and after them.
4. **Genetics Model Revising**: 12 days, with 4 days allotted for each round, including the GCK work and conference, but this is flexible.
5. **Manuscript Preparation**: 5 days, interspersed with other activities over a period of two weeks.

   In addition to the major activities listed above, there are "miscellaneous" activities that occupy some time.

Sue Johnson, when she provided these estimates, emphasized that these are approximate times, and that the schedule is flexible. If she thinks the students are moving along quickly and have finished their useful work on an activity, the time can be shortened. But if students need more time, they get more time.

These five activity groups contain the eleven instructional activities that are analyzed in Sections 3.31-3.35. Before doing this, however, the next two sections examine two important aspects of the course: the classroom atmosphere, and the genetics problems.

3.24: Creating a Classroom Atmosphere

As described earlier, the course developers have tried to create the atmosphere of a "professional scientific community" in the classroom. Many activities, small and large, contribute to this atmosphere.

A. Students as Scientists

During the course the teacher refers to student groups as "research groups," and the classroom conferences are “a gathering of the scientific community.” Early in the course, after hearing brief biographies of prominent scientists (such as Gregor Mendel, Walter Sutton, Thomas Morgan, Barbara McClintock, James Watson, Francis Crick, and Rosalind Franklin), students choose the name of a scientist (it is usually one of these, but not always) to be the name of their research group. Each student receives a personalized badge; for example, Jane Smith's badge might read “Dr. Jane Smith, Senior Researcher, Dept. of Genetics, Harvard University.” Later in the course there is a “Nobel Prize ceremony” — complete with stately music, pictures of the hall where the prize is awarded, scandinavian pastries, and personalized prize medals — to honor the students' achievements. And at the end of the semester, stu-dents “publish” the results of their research in "The Proceedings of Monona Grove Science."

Of course, these symbolic details would be mockingly hollow if there were no substantive reasons for students to feel like scientists. But this is not a difficulty because the most important characteristic of the course, the prominent feature that contributes most to producing an atmosphere of “science in progress,” is the abundance of activities — especially those involving model-revising problem solving — that let students participate in some of the essential experiences of a research scientist.

B. Stories about Science

Throughout the course, students' first-hand experience in “doing science” is supplemented by second-hand experience in science, with stories about the exciting adventures of research scientists, and discussions about the connections between these stories and what students are doing in the course. The science stories include a movie, *The Double Helix,* about the discovery of a structure for DNA. Students also read (or watch) *Alex,* the story of an afflicted child and her family as they struggle to cope with the physical and emotional burdens that accompany the lingering and eventually fatal genetic disease, cystic fibrosis. Later, Sue describes the latest breakthroughs in the battle to counteract the effects of the defective gene that causes cystic fibrosis. The characteristics of these second-hand experiences — and their cognitive and affective functions in the overall context of the course — will be examined in detail in Section 3.47.

C. Metacognitive Reflection

To help students learn more from their experience, the teacher asks them to keep a journal that includes personal reflections about their own experience and the process of science. This *metacognitive reflection* can occur during a problem-solving activity, or after it is finished. In addition to general exhortations to "think about what you are doing, and write in your journal," sometimes there are specific assignments. For example, students could be asked to respond to questions such as "How did you assess the adequacy of the model you constructed?" or, after the detergent box activity and the visit by Mendel, "How was Mendel's model-building process similar to the process used by your research team to construct an explanatory model of the detergent container? (Johnson, 1996, p. 35)" And, to increase their empathy with Alex and her family, students are asked to write from the personal perspective of one of the main characters, expressing their feelings about how the genetic disease has affected their lives.

Students also receive feedback from other students. For example, "The record keeping [in a lab notebook made by each research group] began with the detergent carton activity and was analyzed for completeness by other research groups prior to the three model-revising rounds, thus encouraging more complete accounts of each group's research. (Johnson, 1996, p. 55)"

D. Social-Intellectual Interactions

In a guided inquiry approach such as that used in this genetics course, it is important to aim for an appropriate level of difficulty. If problems are not challenging the students will become bored, and they will learn less than they could. But if problems are too difficult there can be frustration, failure, and a lowering of confidence. One way to adjust the level of difficulty is to aim high, with problems a little more difficult than most students can cope with, and then provide help as needed during the process of problem solving. This type of in-process 'coaching' is incorporated as an integral part of Sue Johnson's course. During each activity, Sue visits each group to see how they are doing, and to ask questions, answer questions, and offer advice. She also provides encouragement and emotional support for groups (and individuals) who are struggling. And, as described earlier, she tries to promote the feeling of a scientific community in the classroom. There are also important student-student interactions, especially between members of a research group who are working together, in cooperative collaboration, in an effort to reach their shared goal of successfully solving problems.

3.25: Genetics Problems in the Classroom

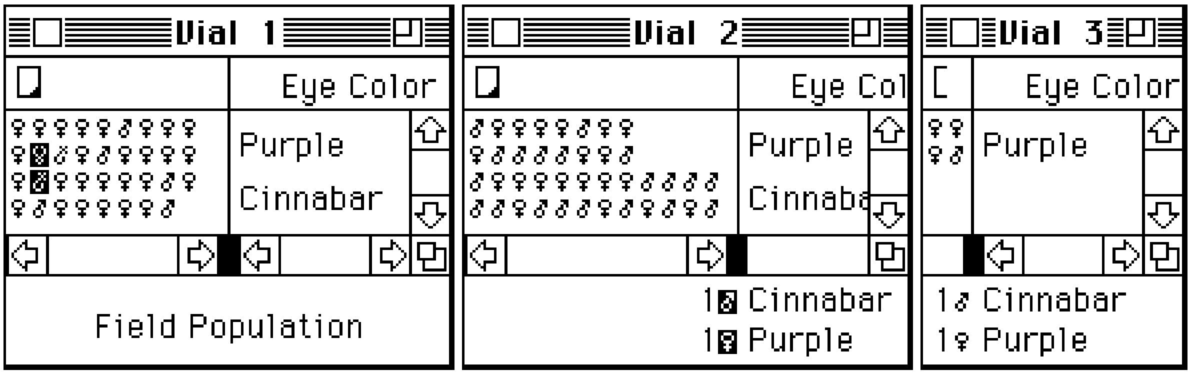
The central activity of the course is model revision using genetics problems generated by GCK. To understand this activity, it is necessary to understand GCK and genetics problems, and why there is a need for model revision. This section describes the GCK problems that are solved by students.

A. Genetics Construction Kit (GCK)

Genetics Construction Kit (GCK) is a computer program (Jungck & Calley, 1984) that generates a simulation of a laboratory (circa 1900-1920) where scientists are solving classical genetics problems. When GCK generates a genetics problem, students are provided with data about one or more *traits* (eye color, eye shape, wing shape,...) for each fruitfly in a *field collection* that is a sample of a larger population; for each fruitfly the sex (male or female) is also specified. For each trait there are two or more *variations*; for example, if the trait is eye color the variations might be eyes whose color is brown, cardinal, plum, ruby, red, or white. To generate more data about their computerized fruitflies, students can select two organisms (one male, one female) from the field collection, and *cross* these two (i.e., mate them) to produce offspring. When a cross is made, GCK generates a statistically realistic *progeny collection,* and shows the trait-variation and sex of each organism in the progeny collection, just as it did with the original field collection. Students can continue doing cross-experiments (by mating organisms from any of the collections, old or new) to gather data.

Figure 7 shows typical data seen by students on their computer screen. First there is a field collection; when the two highlighted two organisms are crossed, the progeny are shown in Vial 2, as are the parents; but when other parents with the same phenotype (but different genotypes) are selected the results are different, as shown in Vial 3 (with its window shrunk so only 4 of the 47 "purples" are shown).

Figure 7: Computer Screen showing a Field Population and the progeny from Two Crosses



In the effect-to-cause problems produced by GCK, students begin with data about observable effects, and then try to propose a theory (by selection or invention) that can explain the cause of these effects. Usually the goal of problem solving is to discover a cause-effect relationship — an *inheritance pattern* theory — between an organism's observable *phenotype* (its total set of traits,[19] each of which can exist in different forms called variations) and the unobservable *genotype* (its total set of *genes,* each of which can exist in different forms called *alleles*) that is causing the phenotype. In GCK the 'inheritance pattern' relationship is totally causal, with the genotype determining the phenotype; in other words, all environ-mental factors (such as what a fly eats, whether it is healthy or ill,...) are ignored by GCK.

GCK lets a teacher set the parameters for the type of problem that she wants GCK to generate. The parameters that can be varied include: the minimum and maximum number of organisms (these can vary from 2 to 999) in a field collection and progeny collection; number of traits (from 1 to 6, selected from 8 possible traits); inheritance pattern; and the types 'analysis information' that are available. For purposes of the MG course, the most important parameter choice is the inheritance pattern that governs the cause-effect relationship between an organism's genotype and phenotype.

---

19. GCK generates a subset of the organism's phenotypic traits, not the "total set" that is the phenotype. But even if students are working with only one trait (of many) and one gene (of many), for linguistic simplicity I will refer to these subsets as the 'phenotype' and 'genotype'.

Even if each of the 8 groups is working on the same type of problem, with the same inheritance pattern, there will be 8 different problems. These are all variations on a basic theme because, for each inheritance pattern, GCK can generate data involving any of the 8 possible traits, with the variations per trait ranging from 2 to 19. Therefore, even if all students are working with the same inheritance pattern, the field collections generated by GCK for different groups will show different traits. And if several groups do have fruitflies with the same trait, the variations will usually differ. For example, even if three groups have data for eye color, the colors for one group may be brown and ruby, while another group has plum and white, and the third group has plum and red.

GCK has other features, but those described above are the main ones used in the MG classroom.

B. A Structured Representation of Mendel's Model

During the "Initial Models" group of activities described above, students are "introduced to a structured way of thinking about Mendel's model." This structured representation is usually called “Mendel's bible” (or “Mendel's Bible” or “the Mendel Bible”) based on the definition of a *bible* as "any book accepted as authoritative or reliable. (The Random House Dictionary, 1980)" A recent version of Mendel's Bible (from Johnson, 1996, p. 36) is shown in Figure 8. This representation contains three major sections: for objects, states, and processes. The ‘objects’ section shows relationships between genes, alleles, traits, and

Figure 8: Mendel's Bible — an external representation for a model of simple dominance, in terms of Objects, States, and Processes.

Simple Dominance Model

Objects in the Simple Dominance Model	
# of Traits under Consideration	1
# of Variations/Trait	2
# of Genes/Trait	1
# of Alleles/Gene in the Population	2
# of Alleles/Gene in an Individual	2
# of Possible Allele Combinations	3

States in the Simple Dominance Model

Allele Combinations: 1,1 or 2,2: Homozygous
1,2: Heterozygous

Genotype/Phenotype Combinations:

1,1	Variation A
1,2	Variation A
2,2	Variation B

Cross Possibilities:

Phenotypes	Genotype	Probable Results
1. Variation A x Variation A	1,1 x 1,1	Variation A
2. Variation A x Variation A	1,1 x 1,2	Variation A
3. Variation A x Variation A	1,2 x 1,2	Variations A & B (3:1)
4. Variation A x Variation B	1,1 x 2,2	Variation A
5. Variation A x Variation B	1,2 x 2,2	Variations A & B (1:1)
6. Variation B x Variation B	2,2 x 2,2	Variation B

Processes in the Simple Dominance Model

Segregation: separation of the members of a pair of alleles such that each gamete produced contains only one member of the pair.

Independent Assortment: segregation of the members of pairs of alleles for two traits, the genes for which are found on separate chromosome pairs. Therefore the segregation of one pair is independent of the segregation of the other pair.

Fertilization: the union of sperm and egg each with a single allele resulting in a zygote with pairs of alleles.

variations, for each individual and for the population.[20] The ‘states’ section shows all possible combinations of alleles, the genotype-to-phenotype mappings, and all possible results of phenotypic matings. The ‘processes’ section describes those aspects of Mendel's model

---

[20]. There is a difference in terminology between a GCK screen (with the first vial labeled "Field *Population*") and the "field *collection*" described above. This change, from population to collection, was suggested by James Stewart because the initial set of fruitflies is a sample from a larger population that exists in the natural environment, and technically a sample should not be called a population. Similarly, the set of progeny that results from a cross is not a ‘population’ in the way this term is defined in population genetics, so this set of flies will be called a "progeny collection."

(segregation and independent assortment) that are most closely associated with meiosis.

Although the core concepts and basic organization have remained the same since the beginning of the course, some details of Mendel's Bible have changed. For example, an earlier version by the course designers (in Hafner, 1991, p. 34) describes the 'processes' in a slightly different way:

> **Segregation**: separates one member of a pair of alleles of a parent during gamete production. Segregation changes the state of an object (an allele pair) from double to single. **Independent Assortment**: assorts allele pairs so that members of the pair for one trait are distributed to gametes independently of the members of pairs for other traits. **Fertilization**: joins gametes that have a single member of each allele pair so that a zygote has two members of a pair.

This earlier version more accurately describes the actual model of Mendel (1865) which did not include any mention of chromosomes.

C. GCK Problems that require Model Revising

The students' initial model of genetics, summarized in the Mendel Bible, is accurate for some genetics systems but not for others. During each round of GCK model revising, one component of Mendel's Bible is revised.

Codominance. In Round 1 there is a change in one mapping for genotype-phenotype causality. With Mendelian *dominance,* two genotypes ('1,1' and '1,2') both produce the same phenotype; because the 1-allele determines the phenotype in the heterozygous '1,2' genotype, the 1-allele is called 'dominant' while the 2-allele is 'recessive'. But with *codominance,* each of the three possible genotypes (11, 12, and 22) produces a different phenotype; in the heterozygous 12, each allele contributes to the phenotype, so neither allele is recessive.

Multiple Alleles. In the Mendelian Model, only 2 alleles exist in the field collection. But in Round 2, there are 3 alleles in the field collection. With 3 alleles, each fruitfly (which still has two alleles) can have 6 possible genotypes: 11, 12, 13, 22, 23, 33. Although many different genotype-phenotype mappings are possible, in Round 2 the GCK parameters are set to produce problems in which the 1-allele is dominant over the 2-allele and 3-allele, and there is a codominant relationship between the 2-allele and 3-allele. Thus, four phenotypic variations are possible: each genotype containing the dominant 1-allele (11, 12, 13) produces the same phenotype, but the other three genotypes (22, 23, and 33) produce three distinct phenotypes.

X-Linkage. In fruitflies, females and males differ in one chromosome; females have two long X-chromosomes, while males have one long X-chromosome and one short Y-chromosome. An X-linked trait is determined by a gene that is located on the X but is missing from the shorter Y; for such a trait, the male has only one allele. If there are two alleles in a collection, a female can have any of three genotypes (11, 12, or 22) but a male can have only two genotypes (1-, or 2-). In GCK the main observable result of X-linkage is that, for some crosses, the distribution of trait-variations will differ for females and males. Before 1994, X-linkage problems always involved dominance. Since 1994, however, two types of problems occur; with problems generated by the GCK file the teacher puts onto the computers used by some groups, in the heterozygous genotype one allele is dominant; but other groups, who are given a GCK file with different parameter settings, get problems in which the two alleles are codominant.

Autosomal Linkage. In the Mendelian Model, traits are independent; in the progeny that result from a cross, variations in one trait are never correlated with variations in another trait. But this ‘independent assortment’ occurs only if the genes that code for two traits are located on different chromosomes. The opposite of independence occurs when two genes are located on the same chromosome, so close together that recombination due to ‘crossing over’ during meiosis is extremely rare. In this situation the progeny that result from mating some parents (but not others) will show a correlation between traits, with some two-trait combinations being more common than predicted by Mendel's Model. If the two genes just described are located on a non-sex chromosome (i.e., they are not on an X or Y) the link between these genes is called an ‘autosomal linkage’.

In Round 3 the use of the last two types of problems (X-linkage and autosomal linkage) has changed over the years. In Round 3, before 1993 the problems were mixed; some groups worked on X-linkage (with dominance) while others got problems with autosomal linkage. In 1993, all groups worked on the same type of problem: X-linkage with dominance. After 1993, some groups got X-linkage with dominance, while others saw X-linkage with codominance. In addition, in all classes autosomal linkage has been given as a challenging "Nobel Problem" to groups who finish early in a round, and for those who want “something extra” to work on.

Each type of problem requires a different type of revision to Mendel's Bible. In Round 1 the major change is that for the heterozygous genotype, instead of dominance (with "1,2 → A") there is codominance (with "1,2 → C"). For Round 2 the major change is that "# of Alleles/Gene in the Population" changes from 2 to 3; another change is from pure dominance (in Mendel's Bible) or pure codominance (in Round 1) to a combination that uses both patterns. With an X-linked trait, for males the "# of Alleles/Gene in an Individual" is only 1. And with autosomal linkage, "independent assortment" does not occur.

3.26: Science Experiences

The ISM framework, which describes the experiences of scientists, was used to produce a list of 17 'science experiences' (split into 7 clusters) at the left border of Table 2. The second column refers to sections in Chapter 2 where each experience is discussed.

Table 2 (is in the "tables.pdf" file)

A brief description of the 7 clusters:

1. The first three science experiences, *PREPARATION, POSING,* and *PROBING,* are part of my '4Ps' extension of the '3Ps' model (by Peterson & Jungck, 1988).

2. The next two experiences, *SELECT an old theory* and *INVENT a new theory,* combine to form one of the four major science-actions of ISM.

3. Two more major science-actions are *DESIGN EXPERIMENT* and *DO EXPERIMENT.*

The fourth major science-action is theory evaluation, which is influenced by the three types of factors in clusters 4-6.

4. The 'empirical factors' that influence theory evaluation involve making *PREDICTIONS,* and estimating degrees of *AGREEMENT* and *CONTRAST,* for both current experiments and *PREVIOUS* experiments.

5. The 'conceptual factors' in evaluation include a theory's *INTERNAL characteristics* and its *EXTERNAL relations with other theories.*

6. The ‘cultural-personal’ factors in evaluation (and in the entire process of science) are summarized as *metaphysical & ideological,* and *psychological, practical, authority.*

7. The result of *EVALUATION* is to retain, revise, or reject; the final P (of the 4Ps) is *PERSUASION.*

Due to the close connection between ISM and the A-and-E grid, whenever a student action is described in terms of a science experience during my analysis, it is being described in terms of ISM.

3.27: Three Stages of Analysis

My analysis of the MG classroom occurred in three stages: a preliminary preparation of the activities-and-experience grid, followed by two phases of analysis. After the planning and goal-setting, I defined the student activities and science experiences that are listed as analytical categories on the horizontal and vertical axes of the A-and-E grid. The first phase of analysis resulted in a description of the science experiences that occur in each of eleven instructional activities. The first phase of analysis served as a foundation for a second phase of analysis that examined the functional relationships within and between activities, in an effort to understand and describe the structure of instruction in the course.

3.28: Sources of Information for the Analysis

Information used in characterizing the MG course came from: published sources that contain descriptions of the course curriculum, its philosophical foundation, and analyses of students' problem-solving behavior; interviews with two of the three course developers (Sue Johnson and James Stewart); and nine weeks of my own observations of the MG classroom.

A. Methods for the Central Activity

In Section 3.34_A the central activity of the course, genetics model-revising problem solving, is examined in careful detail. This examination took advantage of the currently existing "analyses of students' problem-solving behavior" by sifting through 1100 kilobytes of annotated transcripts and analytical conclusions in four doctoral dissertations (described in Appendix B20, by Finkel, 1993; Wynne, 1995; Lemberger, 1995; Johnson, 1996) written by researchers who have examined the classroom. The sifting that eventually resulted in Section 3.34_A occurred in four phases.

In the first phase I wrote a detailed description of students' problem-solving behaviors in a long section (Appendix B22) that expresses my own ideas and also incorporates insights, empirical evidence, and illustrative examples from the four previous researchers. This description, although fairly comprehensive, does not attempt to cover every detail; instead, I focused on the empirical observations and theoretical claims that seemed the most interesting and important. In the second phase I reviewed Appendix B22 and searched for ‘science experiences’; each experience is indicated, in B22, by bold-faced fonts, and is accompanied by a reference to the section where this experience is discussed in the ISM elaboration. In the third phase I summarized these experiences (using the 17-category format of the activities-and-experience grid) in Appendix B21, which provides an overview of B22. In the fourth phase, ideas from B21 and B22 were combined and revised to form Section 3.34_A, which is intended to be self-contained, independent from B21-22.

The description of student experiences in Section 3.34_A is based on empirical evidence (from my own observations and from the four researchers) about what students actually do in the classroom. But this evidence has passed through several stages of interpretation. Initially, student groups were audiotaped as they worked on GCK problems. Next, these tapes were transcribed by a researcher. Then these transcripts were studied by the researcher, and only those conversations and actions that were judged to be the most interesting and important appear in each dissertation. Finally, I did my own sifting and interpreting.

Despite these levels of processing and interpretation, there are two reasons to think that my description is built on a fairly solid empirical foundation. First, I have no reason to suspect that there has been any significant loss of empirical integrity in moving from the raw data (on the audiotapes) to the four dissertations I used as sources. Of course, some data has been lost due to selectivity by the researchers, but I am assuming that the student behavior described in the dissertations is a fairly representative sampling, or (in a weaker claim) that it is suitable for purposes of instructional analysis. Second, the empirical basis for what I have written is under-represented by the sources I have cited, so the support is actually stronger than it might seem. For example, in Appendix B22 most of my claims about science experiences are supported by only one illustrative example, but in many cases other examples also exist in the transcripts; I just chose to quote or cite the example that I thought was the best illustration. Similarly, I often cite an interpretive claim made by one researcher, rather than citing all of the similar claims that have been made by other researchers. And in moving from B22 to 3.34_A the empirical support appears to weaken, because in 3.34_A I have omitted most references (that are cited in B22) for the sake of easier readability. The interested reader can, of course, proceed from Section 3.34_A to B22, and onward to the references cited there. In fact, I recommend this, since I think the "story" in Appendix B22 is one of the more interesting parts of my dissertation project.

B. Methods for Other Activities

For 10 of the 11 activities in Sections 3.31-3.35, there is no direct empirical evidence such as that described above. Because of this, the claims made for these activities is weaker than the claims for GCK model revising. The following paragraph describes how the difference in data sources affected my analysis of the Black Box activity. Similar differences occurred for the rest of the ten "other activities."

Although the process of model development is similar whether students are studying a black box (in Activity #1) or fruitflies (in Activity #4), the latter process has been observed more carefully, and has been interpreted in greater depth, by researchers who have studied the MG classroom. Therefore, due to a relative shortage of empirical data (and theoretical conclusions) about problem-solving strategies used during the Black Box Activity, the analysis in this section makes no claim to be a definitive portrayal. Instead, it is an estimate, based on four main sources: my observations of students (in Sue's classroom); my own experience (in a seminar with a group of fellow graduate students) in building and revising models for the special carton; conversations with other educators, especially Sue Johnson, about black box activities; and the "probable action" estimates that are described in the next two paragraphs.

One approach to characterizing student experiences makes a distinction between student activities and student actions, as described in Section 3.22. By asking and answering the question, "If students make a sincere effort to accomplish what they are asked to do, what actions will they probably do?", it is possible to make an educated estimate, even without observing student behavior, of the actions that will probably occur (by necessity or choice) in each activity. This "probable action" approach cannot be defended by rigorous logic, but nevertheless it is often a useful heuristic device. An educated estimate of probable actions is a major source for the 'science experiences' described in Sections 3.31-3.35, for all eleven activities. For each activity, this estimate is supplemented by empirical observations, both first-hand (made by myself in the classroom) and second-hand (from written and oral sources), but this supplementation is much greater for the GCK model-revising activity, compared with the other ten activities.

3.3: The First Phase of Analysis — Student Experiences in Each Activity

Using the methods described in Section 3.2, Section 3.3 analyzes the science experiences that occur during each of eleven instructional activities in the MG course.

3.31: Activity Group #1 — Black Box Model Revising

In this activity, the teacher asks each research group to examine a 'black box' object and to develop a model that explains how the object works. Sometimes the black box is a detergent carton, as described earlier, but other objects can be used if they behave in puzzling ways and if students will be challenged to think creatively and critically in their effort to develop satisfactory models to explain the observed behavior.

A: Developing (building and revising) Models

The following activity-and-experience analysis begins with a description for each cell in the first column of cells (entitled "3.31_A") of Table 3 on page 230.* The 17 categories below are numbered from 1a to 7b, corresponding to the numbers used in Table 2.

* **Tables 3-7** are in the "tables.pdf" file

For simplicity, in this section I will assume that a specific type of black box, the special carton, is being studied. This type of black box has been chosen for two reasons: compared with the other black boxes, I am more familiar with the carton; and it was used in two of the three courses that were included in my detailed examination of classroom transcripts.

In the remainder of Chapter 3 the term 'model' will be used as a synonym for 'theory', and these two terms will be used interchangeably.

1a. Preparation? For each activity, there are two ways to view preparation: backward-reaching (what already has been done to prepare students for this activity) and forward-reaching (how this activity helps students prepare for future activities).

Backward-reaching preparation? Students do not explicitly prepare for this activity; they do not learn any initial theory about the structure of detergent cartons (content) or how to do experiments with cartons (process). Therefore, a "NO" is in the top half of this cell. But there is also a "?" because students, in their everyday living and in school, by reading and listening and doing, have learned knowledge involving content (about "how things work") and process (about "how to ask questions and search for answers") that, even though it was not originally learned for this purpose, can transfer to the black box activity.

Forward-reaching preparation? The main function of the black box activity is to serve as a preparation for future activities. The teacher explains, "This activity is done to have the students experience the process of model building and the subsequent process of model revising. (Johnson, 1996, p. 34)" Since analogous procedures are used in developing models for the black box and for GCK problems, when students begin the central activity in the course (GCK model revising) they already have valuable model-developing experience due to their work with the black box.

1b. Posing? There are two main aspects of problem posing: the choice of an *area* to study, and the choice of *constraints* that define a problem and its potentially acceptable solutions. Therefore, in the far-left column of Table 3 there are 'area posing' and 'constraint posing' definitions on the top and bottom lines, respectively. In some activities in other courses, such as making decisions about a science project or a term paper, students have a significant amount of freedom to choose among many available options for "an area to study, and questions to ask about it." By contrast, for the Black Box activity in this classroom the teacher defines the problem area and the general goal when she chooses a type of black box (whether it is a detergent carton, or a box with dowels, or something else) and asks students to answer the question, "How does it work?" Because students do not choose the area being studied, there is a "NO" in the top half of the cell. But there is a "yes" in the bottom half because, although the teacher's request to construct "an adequate explanation for how it works" defines the basic goal, students do have a limited role in determining the constraints that define an acceptable solution. The question of students' involvement in problem posing is discussed more fully in Section 3.44.

1c. Probing? In my opinion, probing is the most significant science experience in the course, thus the "YES!" in this cell. During their pursuit of a solution, students must *invent* options for effective probing actions (to observe and interpret the system being studied), *evaluate* these options to decide what to do next, and then *execute* these actions. Probing actions are done for the purpose of pursuing a problem solution, and can include any of the actions in ISM. Some probing actions (design experiment, do experiment and make observations) are done primarily to improve *observation* knowledge, and some (theory selection and invention, empirical evaluation, conceptual evaluation, and drawing conclusions) involve the *interpretation* of observations. During the black box activity, students execute (after invention and evaluation) a wide variety of observation-and-invention actions.

2a. Selection? The special detergent carton looks like an ordinary milk carton. Unless students have led a sheltered life they have seen and used milk cartons before, so they have an initial theory about cartons. Based on observations about the special carton, however, especially its ability to dispense a fixed amount of liquid, their initial theory — that a carton is just an empty volume with no special mechanism inside — will be judged inadequate. But this simple theory about ordinary cartons can serve as a starting point for inventing more complex theories about the special carton. In addition, individual students will have their own general theories about how *other* things work, and these theories, when applied by analogy to the black box situation, may be useful.

2b. Invention? Yes, retroductively inventing a new theory is the main goal. Because the problem is difficult enough that students' earliest invented theories are never totally adequate, there is always a need to revise these new theories. Students realize this, so from the beginning they engage in both model building and model revising.

3a-3b. Experimental Design, Experiments and Observations? In order to do retroductive invention, students need observations. Each research group has its own carton, so students soon supplement their ‘field study’ observations, made by watching what happens when Sue tilts a carton and then pours, with their own ‘controlled experiment’ observations. Students can tilt and pour, then tilt (in a different way) and pour, they can try to look inside the carton, tap it on the outside, check for its balance point, and do almost anything they want — Sue's only rule is, “Don't tear the carton apart!” — in an effort to improve their database of observations about the carton. Although initially students have general theories (about gravity and other laws of nature, and about cartons), early in the process of probing the students usually just “do things to see what will happen.” It is in later stages, after specific theories about the carton have been formulated, that theory can have a more direct effect on the design of experiments.

4a-4d. Empirical Evaluation? Students are not required to predict what will happen before they do an experiment, and they usually don't make pre-experiment predictions. But theory-based predictions are more common after an experiment, when they are done for the purpose of evaluating an old theory or inventing a new theory. In their empirical evaluations, students mainly check for degree of agreement between observations and the predictions made by one theory. But if a group has proposed two or more theories, there may be a ‘crucial experiment’ with predictive contrast because competitive theories make different predictions for the same experimental system. As they continue to run experiments and collect data, students consider results from their most recent experiment and also from previous experiments.

5a-5b. Conceptual Evaluation? In addition to empirical adequacy, students also expect their ‘specific theory’ about the carton to be consistent with their own ‘general theories’ about how the world works. In evaluating conceptual adequacy, students can ask whether a particular model is physically possible, and is consistent with known physical laws such as gravity and with the known properties of liquids and solid walls. Other possible conceptual questions involve engineering and business decisions, such as whether it would be practical, physically and economically, to massproduce the type of carton postulated by a model.

6a. Metaphysics and Ideology? The most common metaphysical theory is an assumption of ‘consistency’ — that if the same experiment is repeated, the same results should occur. If there is any ideology attached to the cartons, it is too subtle for me to recognize.

6b. Social factors? I find it useful to distinguish between process and content, in order to understand the ways in which science process and content are (and are not) affected by cultural-personal factors. Although ‘authority’ can exert a significant influence on students, I doubt that the other cultural-personal factors (psychological, practical, metaphysics, ideology) have much effect on the *content* of the models eventually accepted by students. On the other hand, the *process* of model construction is influenced by interactions of students with the teacher and with other students. The major student-student interactions occur while inventing and evaluating actions and theories. But there are also small social decisions such as dividing tasks within a group — by deciding who designs experiments and runs them, and who does what with the lab notebook (who takes notes on experiments and observations, who draws pictures of models, and so on).

7a-7c. Conclusion and Persuasion? During the process of model building, model evaluation occurs often, and so do decisions about whether various models should be retained, revised, rejected, or “put on hold” by temporarily delaying judgment. When students are evaluating, as individuals or as a group, persuasion occurs at three levels: there is a persuasion of self, of other students in the research group, and of the group as a whole. For each model constructed by a group the result of evaluation is an estimate of ‘status’, with a degree of confidence rather than certainty, but as a group the students eventually must make binary “yes or no” decisions about which model(s) they will show during the next day's conference. A group can also prepare a ‘persuasion strategy’ — including experiments and external representations (pictures of a model, plus verbal explanations and arguments) — that will be used for the conference.

B: A Student Conference

In order to more efficiently describe student experiences in the second activity, I will focus on the differences between what occurs during the first activity (model construction) and second activity (the conference). Also, my descriptions will continue to follow roughly the order in the A-and-E grid of Table 3, beginning at the top and working toward the bottom.

Looking backward from the second day, the first day of model development has helped students prepare for the conference. And looking forward, the second day is a preparation for the following day's continuing model development, because during the conference students have an opportunity to "listen to ideas of other groups" and thereby gain new perspectives on the composition-and-operation of the black box. More important, however, the conference and the two days of model revising are a valuable preparation for the remainder of the course.

In an effort to pursue their goal of effective persuasion, a group will invent, evaluate and execute pursuit-actions, choosing from a wide range of possible contents and styles. But now the objective shifts from heuristic action (for the purpose of learning) to demonstrative action (for the purpose of persuasion).

A group's own ‘old model’, constructed the previous day, is the focus of a group's conference presentation. But while listening to the ideas of other groups, or while presenting their own model, students can decide (individually or as a group) that their model will be improved by a quick, improvised revision.

A group, to illustrate and support a model it is presenting, will choose the most impressive ‘demonstration experiments’ from the previous day. And other students — in an effort to challenge the model, and maybe to thereby falsify it — can also design an experiment. This could be an old experiment (already done by the challenger's group) or a new experiment. If a group is challenged to do a certain experiment, the teacher asks the group to first predict what will happen. After this the experiment is performed, observations are made, and everyone can draw conclusions about whether or not the results provide empirical support for the model. There can also be a conceptual challenge if there is suspicion that a model is inconsistent with accepted theories.

In the conference many alternative models are presented, and students can compare the ‘overall appeal’ for each of these competitive theories. During a presentation there is externally-oriented persuasion, done for the purpose of convincing others that the group's own model is satisfactory. But there is also internally-oriented persuasion, less obvious but nevertheless real and important, where members of a group, either privately or for others to hear, ask themselves whether their model's status has increased or decreased, and whether they still consider the model to be satisfactory.

In the conference some new social factors emerge. Instead of informal conversation in a small group, there is public speaking with a large audience. Because the teacher encourages active dialogue between the presenters and audience, this often leads to direct interactions between students from different groups. And within each group, new tasks and roles must be decided and delegated; for example, during the presentation will there be one main spokesperson (and if so, who) or several (and if so, who will do what, when, and how will group members interact).

C: Revising Models

In the Table 3 summary, most of the boxes in the "3.31_C" column have a simple "yes" or "no" to indicate that the third day (model revising) is similar to the first day (model building). The main difference is that each group begins the third day with at least one 'old theory'; this is probably a model the group itself has constructed, but it could be a model they learned about (and were impressed with) from another group during the conference. Thus, instead of constructing a totally new model, students can revise an old model. By the end of the day the group will have more experiments and data to work with, and will be expected to decide on a "final" model that has high relative status; the intrinsic status of the final model may be either high or low.

3.32: Activity-Group #2 — Genetics Phenomena

This group consists of three activities — cookie analogy, human variations, and human pedigrees — whose purpose is to introduce students to "the types of phenomena, such as similarity, diversity, and continuity, for which geneticists attempt to build models. (Johnson, 1996, p. 34)"

A: The Cookie Analogy

As a homework assignment, students bake different types of cookies by using different recipes and their own variations of these recipes. Then in class, in their research groups, they examine the cookies and discuss some thought-stimulating questions — for example, about the relationship between the recipes (the genotypes) and the resulting cookies (the phenotypes). The educational potential of this activity is described by the teacher:

> The questions can be used to introduce and/or support all of the following concepts: gene, allele, heredity, environment, genotype, phenotype, adaptation, evolution, adaptation, evolution, continuity, diversity, mutation, mutagen, teratogen, taxonomy, and classification. For example, margarine, butter and shortening can be used as examples of *alleles* (forms of a gene). The value of the analogy becomes more evident as it is applied to help the student construct meanings for concepts that are unfamiliar to their daily experience and therefore difficult to understand. At many times in class discussions — days or weeks following this activity — the familiar, concrete example can support the unfamiliar, difficult concept. (Johnson, 1991, p. 3)

As indicated in the first cell of Table 4, in this activity students "learn concepts and terms" that will help them later in the course. Another important experience is the social bonding that occurs, in research groups and in the class as a whole, when students share a good time (observing and discussing, plus milk and cookies!) with each other.

B: Human Variations and Human Pedigrees

In the 'human variations' activity, students study similarity and diversity at the level of individuals. To answer the important questions, "In what ways are individuals within a species similar, and in what ways are they different?", students check the phenotypes of themselves and their classmates for certain traits. While observing self and others, and discussing "who has which variations of the trait," there are more opportunities for social bonding.

In the 'human pedigrees' activity, which is done mainly as a homework assignment, students explore similarity and diversity at the level of a family by examining family trees and searching for patterns that involve the inheritance (or non-inheritance) of traits such as blue eyes, cystic fibrosis, and color-blindness.

In Table 4 most of the cells are blank because there are opportunities for only a few types of experience. Students can learn genetics concepts and the associated terminology, and they may be stimulated to do 'question posing' about the phenomena they are observing. In one activity students make observations, in the other they use second-hand data provided for them. When students analyze human pedigrees to search for phenomena that are interesting or puzzling, they can practice and improve their pattern recognition — the art of finding patterns in data — which is a cognitive skill that often plays a valuable role in research science.

3.33: Activity Group #3 — Initial Models

In this group of activities, students learn an initial model of classical genetics. This model contains two sub-models: dominance and meiosis. In addition, students learn how to use a computer program (GCK) that lets them become more familiar with their initial models of genetics. For purposes of discussing activities-and-experiences, this section (and Table 5) will be split into three parts: developing a Mendelian model, developing a model of meiosis, and using these models with GCK.

A: Developing a Mendelian Model

To construct the model of Mendel, students begin by reading Mendel's classic paper (1865); then they are visited by a person dressed as Gregor Mendel. 'Mendel' brings pea flowers and three generations of peas, which students classify and count, in a reconstruction of the process that led Mendel to develop his model of simple dominance.

As summarized in the first cell of Table 5, the main function of this activity, which results in learning an initial model of genetics, is to serve as 'forward-reaching preparation' for future activities that require revising this initial model. Also, within this activity there is preparation on a smaller scale; reading a slightly edited and annotated version of the paper by Mendel (1865) serves as a preparation for the model-building that occurs during Mendel's visit.[21]

In this activity the agenda has been set by the teacher, so there is no area-posing by students. And most pursuit-actions are planned by the teacher; students are not responsible for deciding "what to do next."

Most students in the course have taken a genetics unit in which they spent several days learning

---

[21]. As discussed in Section 3.21, the definition of 'activity' is flexible. In this section, for the purpose of analysis the reading and visit are being treated as components of one activity, but near the end of Section 3.23 these components are listed as two activities.

the inheritance pattern of simple dominance. But during Mendel's visit students are expected to assume that Mendel's Model is not "an old theory they already know" that should be accepted due to the authority of a teacher or textbook. Instead, it is a theory that is being developed by the students, with guidance from the teacher, based on observations and logic.

Students do not design an experiment but they do experience, vicariously, the expert high-quality designing done by Mendel. The data source (the collection of peas) is provided for them, but students do play an active role in the data-gathering process as they classify peas — by determining for each pea the variations for two traits (is the pea round or wrinkled, yellow or green) — and then count the number of peas in each of four categories: round-yellow, round-green, wrinkled-yellow, wrinkled-green.

After the data gathering there is a period of retroductive theory invention and evaluation, based mainly on empirical agreement. Punnett Squares are used to generate Mendelian predictions that can be compared with the data. Empirical evaluation is not based only on the peas counted by students, because Mendel's other experiments are also discussed. Models that postulate a 'blending' of inherited characteristics are considered, but these are judged to be empirically inadequate because in situations with predictive contrast (relative to Mendel's model) these models make incorrect predictions. Mendel's model is examined for internal logical structure, especially when they — the students, Mendel, and teacher — formulate a structured representation (Mendel's Bible) in terms of objects, states, and processes. Basically, the Mendelian model combines the concept of dominance with descriptions of meiotic patterns (such as segregation and independent assortment), but without postulating meiosis as an explanation for these patterns. At this stage in the course (when knowledge is limited to what Mendel knew in 1865), before the class has developed a model of meiosis (which is knowledge from around 1903), "external relations with other theories" is not a major consideration.

For all students, the conclusion of evaluation is to accept the model of simple dominance they have developed. Why do students accept this model? Partly it is due to a logical evaluation of empirical evidence (the peas), combined with the students' assumption that this evidence is legitimate — i.e., that the peas accurately represent what really occurs in nature. But there are other important reasons for acceptance. First, there is powerful persuasion due to the 'authority' that comes from the teacher, from the prestige attached to Mendel, and from what students know is taught in textbooks. Second, a practical reason for at least an appearance of acceptance is that students are enrolled in a course, and the teacher (who wants them to accept the theory) will determine their course grade. Third, another practical reason is that since the model-revising activities later in the course depend on having an initial model to serve as a foundation for revision, if students don't accept dominance they will have no starting point for these future activities.

B: Developing a Model of Meiosis

The process for building a model of meiosis is similar to the process, described above, for developing a Mendelian model. Again, students read an edited, annotated paper; this time the paper is by Sutton (1903), who wrote about chromosomes and heredity. And again students develop, in a process guided by the teacher, a model whose purpose is to adequately explain observed phenomena.

As with the Mendelian model, in this course the main function of the meiotic model is to serve as a preparation for future activities in which the model will be revised. To produce knowledge gaps that let students solve ‘mystery stories’ later, a limited model of meiosis — a model that is incomplete without being incorrect — is constructed at this early stage in the course.

Students begin their model building with an initial model of mitosis that can adequately explain biological growth and asexual reproduction. But this model is not consistent with the observation that in sexual reproduction the parents and offspring are not identical, and that offspring usually differ from each other. Therefore, the mitotic model must be revised to construct a model that provides a more adequate explanation for these phenomena, and for other observations (such as the "genetic phenomena" encountered earlier in Activity Group 2) that involve inherited similarity-and-diversity. The crucial anomalous observations — i.e., those that are incompatible with mitosis, and must be accommodated by a revised model — act as constraints on a problem solution, and when students recognize these constraints (and their role in the process of problem solving) this can be viewed as constraint-posing.

The anomalous observations come from field studies of naturally occurring phenomena, not from controlled experiments designed by students.

After a meiotic model has been developed by using retroductive logic (and prior knowledge of meiosis) the predictions of this model can be checked for agreement with all known data. Students begin with a model of mitosis, and use predictive contrast (i.e., relative degree-of-agreement) to decide that meiosis is a superior model for explaining sexual reproduction.

As usual, students can check whether their model is internally and externally consistent. A model of meiosis should be consistent with the Mendelian model developed in the preceding activity, if both models are valid. In fact, both models are supported when meiosis (based on chromosomes) provides an explanatory mechanism for Mendel's descriptive postulates (which he made with no knowledge of chromosomes, by recognizing empirical patterns) for segregation and independent assortment. If the Mendelian and meiotic models are viewed as two models with shared components, students will ask whether there is external consistency. But if there is one model, which is a Mendel/meiotic synthesis of the concepts of dominance and meiosis, the question is about internal consistency.

To develop a model for meiosis, students initially work inside their own research groups, and there can be pursuit-decisions about which observations to focus on, and how to approach the task of interpretation. Later, the whole class works together to develop a single model of meiosis. As with the dominance model, students accept the meiotic model due to a combination of empirical factors, practical factors, and authority. But compared with dominance, in the evaluation of meiosis there is an increased reliance on conceptual factors because meiosis can be checked for agreement with the existing Mendelian Model that was developed in the preceding activity.

C: GCK Problems without Model Revising

In this section the order of analysis (dominance, meiosis, GCK) is not chronological. In the course, the actual order of activities (dominance, GCK, meiosis, GCK) has students working with GCK both before and after they develop a model for meiosis. During this work there are several types of GCK activities (described in Section 3.23) but the differences between these activities are minor, and the following analysis will treat all of the initial GCK work as a single type of activity.

Forward-looking preparation: The main purpose of this GCK practice is to let students become more familiar with their initial models, and with the Macintosh computer and GCK program, as a preparation for the upcoming model-revising activities. To help them learn the model of Mendel more thoroughly, Sue asks students to think about their experiments and observations in terms of the objects, states and processes in their Mendel Bible. Later, during a GCK-based exam, students are asked questions such as whether a particular trait-variation is dominant or recessive, what the genotype of a certain parent could be, how to support these answers with evidence, and how to explain their experimental results in terms of meiosis.

As usual, the teacher decides the types of problems (in this case by setting the parameters in GCK) so there is no area-posing by students.

During this ‘initial models’ phase the teacher assigns only problems in which GCK generates data that is consistent with the students' single existing model. Because at any given time the students have only one genetics model (first it is Mendel's model, and then Mendel-and-meiosis), selection is easy. And because all data is consistent with this model, there is no need to revise it and thus invent a new model. However, there can be inventive retroduction regarding the experimental systems — for example, by constructing a theory about which trait-variation is dominant, or about the genotype of a certain fruitfly.

When students use GCK they design their own experiments, but their decisions are limited to selecting which parents to mate. They do not, for example, have the option of doing biochemical experiments on the GCK organisms. When students make a cross, the data for progeny is automatically provided by GCK, but students can choose which aspects of this data they will mentally ‘observe’; and they can choose how to cognitively process the data by searching it for patterns, by comparing it with predictions, and so on. To check for degree-of-agreement, students compare observations (from any experiment, current or previous) with predictions that are made either before or after the data is generated. Since there are no alternative models, there is no need to check for degree-of-contrast.

By the time students begin using GCK, their existing models (Mendelian and meiotic) have already been checked for conceptual consistency, so there is no need to do this again. But students can deepen their understanding of the internal structures of dominance and meiosis, and of the relationships between dominance and meiosis.

As discussed previously, students usually adopt a metaphysical assumption of consistency, that identical systems should produce consistent results. But during a computer simulation of nature this is actually an assumption about the computer program, not about nature. When students assume consistency, they assume that GCK is generating data according to some 'rule governed' logic, and that these rules will not change in the middle of a problem. In classical genetics, however, even though two experimental systems appear to be visually identical if we look only at the surface level of the phenotypes that are observable in GCK, different results can occur if the systems are not really identical at the deeper level of the genotypes that are unobservable in GCK. Students learn this during their GCK practice, and they learn that an assumption of 'empirical consistency' should be interpreted with conceptual understanding and logical sophistication. In addition, a theory often predicts some statistical variability, so an expectation of "consistent results" should be interpreted statistically, not literally.

Persuasion? Even though there is no need to argue about the adequacy of the genetics model being used, members of a research group can still discuss how to apply the model, and they can debate the characteristics of experimental systems. There will be 'pursuit action' discussions about how to use the GCK program, and about which experiments should be done and for what reasons. There also are social negotiations regarding tasks, such as who operates the mouse (and thus the program), who designs experiments, who takes notes on the GCK data, who processes the data and makes predictions, or who asks questions (and answers questions) when the teacher stops to talk.

3.34: Activity Group #4 — Genetics Model Revising

During this group of activities the students work with GCK problems in which their existing models are not adequate for explaining the observations. Before these rounds, students are told that they will need to revise their initial models, and that there are no limits (set by the teacher) on the parts of the models that can be changed.

In Activity Group #4 there are three rounds of problem solving: in Round 1 the students' GCK work is followed by a conference to discuss the results of their work; then comes Round 2 and a conference, and Round 3 with its conference.

A: GCK Problems that require Model Revising

This activity is worthy of especially close study, due to its central role in the course, and because it is the closest that students come to being engaged in a type of work done in actual research science. Therefore, it is described in more detail than the other activities. And if the reader wants to know more about what students do when they revise models, Appendix B22 tells the story in greater depth. The analytical methods used for this section are described in Section 3.28_A.

1a. Preparation: This activity is the "main event" of the course, the focal point for action. Looking backward, all preceding activities (black box, genetic phenomena, initial models) are a preparation for it. And looking forward, all subsequent activities (student conferences, manuscript preparation) are a continuation of it.

Preparation is one of the most important factors in producing success in model revising. If essential content-knowledge or process-knowledge is not cognitively available for students (because they did not learn, or do not remember), problem-solving effectiveness will decrease. But even if students do not begin Round 1 with a solid foundation of content-knowledge and process-knowledge, their knowledge of concepts and procedures can improve. What is learned early in Round 1 will help students later in the round. And what is learned during Round 1 (including its conference) will help at the beginning of Round 2, and so on.

1b. Posing: As usual in this course, the teacher chooses the problems (in order to achieve an appropriate level of problem difficulty), so students do not experience any area-posing.

But students do participate in three types of constraint-posing, as discussed in Section 3.44_B. One type of constraint-posing occurs when an anomaly has been recognized by students, and there is an expectation that the anomaly should be resolved. This expectation about anomaly resolution is the most important constraint on what will be accepted as a satisfactory solution.

1c. Probing: Students are constantly making decisions about probing actions — about how to use time, and "what to do next" in pursuit of a solution. These decisions result in the invention, evaluation, and execution of actions for 'observing and interpreting' the systems being studied. In addition to individual actions, there can be mega-actions that involve several component actions, such as 'testing' a theory by designing experiments, doing experiments, and evaluating theories for either pursuit or acceptance.

2a. Selection: At the beginning of each GCK problem, students use their existing genetics models to interpret experiments and data, and to recognize ways in which the data is anomalous. At the start of Round 1 only the initial Mendelian Model (along with its meiotic mechanism) is available. During Round 1 this model is revised to construct a new model. At the beginning of Round 2 students can choose from two models — their initial model, and their newly invented model — so at this point (but not before, if selection implies a need to decide between options) they can select a model from their two options.

In normal research science, scientists usually do not expect their most important existing 'domain theories' to be revised. But in this classroom the teacher tells students that they will be revising their existing models in each round, so students "expect to be surprised" and the psychology of selection-versus-invention is different than in most research science.

2b. Invention: In each round the previously existing models will not be adequate, due to the problems selected by the teacher, so the invention of a new model is required. A simplified summary of the overall problem-solving action is that students "find what is wrong and fix it"; more specifically, they "recognize an anomaly and resolve it." The main strategy for anomaly resolution, which the teacher encourages and the students actually use, is to revise parts of Mendel's Bible, which — by serving first as a standard for defining expectations and *recognizing* anomalies, and second as a template for inventing new theories that can *resolve* these anomalies — facilitates the two main steps in model-revising problem solving. In fact, the course designers have intentionally structured Mendel's Bible in a way that facilitates the process of analysis-and-revision by allowing the localization of anomaly to a specific component of the model, which can then be revised to produce a new model that resolves the anomaly.

Even if several anomalies have been recognized, students usually choose one anomaly to serve as a focal point for the retroductive invention of a revised model that can resolve this anomaly. When a new model has passed this first stage of evaluation, it is ready for more thorough testing, to check whether it can explain all available data, and whether it is conceptually satisfactory.

So far, the discussion has focused on inventing a domain-theory (such as a revised version of the initial Mendelian Model), but in GCK problem solving it is often just as important to invent system-theories that accurately characterize the experimental systems. One reason for this is that the logic used in genetics is much easier when the genotype of one or more fruitflies is known, because this knowledge restricts the range of cross-result possibilities that must be considered. With this simplification it is easier to make deductive predictions or retroductive speculations about a model or system, and logic is more likely to lead to a definite conclusion instead of a deferred judgment that requires patience and further experimentation.

3a. Experimental Design: Because genetics logic is so much simpler and more powerful when students know the genotype of a parent — such as whether it is homozygous or heterozygous, and containing which allele(s) — strategies for generating this knowledge are very useful. With a little experience, students can develop their own observation-and-interpretation strategies for determining genotypes, and their own efficient algorithms. These strategies and algorithms, oriented toward the specific goal of developing a system-theory, can guide the design of experiments. Experimenting can also be guided by a domain-theory, such as a new model the students have developed, but this is less common.

The most common goal-oriented strategy for designing experiments has a general goal — to gather lots of data, by doing all possible experiments, to use for the retroductive invention of models and system-theories. This type of "do it all" strategy is common because it is useful to have lots of data, and because GCK experiments are quick-and-easy, costing the student very little time and no money. Due to this low cost, there is not much incentive to do 'thought experiments' for the purpose of designing physical experiments, in contrast to actual research science where a physical experiment often requires a significant investment of time and money.

The key to GCK model revising is anomaly recognition, which often (but not always) requires the generation of experimental data. After at least one anomaly has been recognized, however, a group must decide whether to invest more time in observation, or shift to interpretation. Sometimes gathering more data leads to a pattern recognition that inspires an explanatory model. Or it may be more useful to stop experimenting in order to focus on interpreting the existing data.

As usual, there is a delicate balance between investing too much time (in either observation or interpretation) and not enough time. The characteristics of an effective balance vary from one situation to another, and there is no easy algorithm for deciding when to shift from data gathering to model development, or vice versa.

3b. Do Experiments and Make Observations: Although students never do physical experiments, GCK generates data as if an experiment had been done.

Sometimes alert observation is necessary, especially if, in a new situation, an important characteristic of the data has not been important in the past. For example, two principles that are essential in problems with X-linkage and autosomal linkage — that ratios can differ with sex, and that two traits can be correlated — have not been important in the students' previous experience with Mendel's Model, or in Rounds 1 and 2. Therefore, students must somehow learn to pay careful attention to these newly important details. Even though the relevant information (about sex differences or trait correlations) is available on the computer screen, unless it is “mentally observed” this information does not become “collected data” that is available for interpretive processing.

Sometimes the new characteristic makes its presence known in a way that is difficult to ignore. For example, in an ‘X-linkage with codominance’ problem, when a group tries the “do it all” strategy of experimentation, they discover that for one of the three variations there are no males. In this case there is serendipity, because an experiment that was designed for one purpose becomes useful for a different reason. But with autosomal linkage the students must be very observant, just to recognize an anomaly. The most distinctive and informative cross — of a dually heterozygous parent with a dually recessive parent — is quite distinctive in producing an unexpected result (if one knows what to look for) but is a difficult experiment to intentionally design until after a model of autosomal linkage already has been invented and understood. Even then, setting up this crucial experiment requires preliminary experimenting, guided by careful logic, to find parents with the desired genotypes. And if a group gets lucky and does this experiment[22] by random chance, they

---

22. Other experiments also produce anomalous results, but this particular combination of parents produces the most obvious and easily interpreted anomaly.

still must know "what to look for" if they are to take advantage of the opportunity.

4a. Predictions: Students can make a prediction by using deductive logic (based on a theory that is either descriptive or explanatory) or inductive logic (based on a descriptive theory, or on uninterpreted experience). An example of an "inductive prediction without a theory" occurs at the beginning of Round 1 when students recognize that there are 3 variations in the initial field collection, instead of the 2 variations that have always been present in the past. Eventually students can use Mendel's Model to deductively explain *why* there should be only two variations, but initially the anomaly recognition seems to be mainly due to simply recognizing a departure from previous experience.

Predictions are usually made after observations. With this timing, students are checking to see if a model explains old data, not asking whether it predicts new data.

When students test a newly proposed model, making predictions based on the new model is one of the most important steps, and also one of the most difficult.

4b. Agreement: Degree of agreement is used to recognize anomalies, and to retroductively invent a model or system-theory, and it seems to be the most important criterion for evaluating a model.

For many GCK experiments, accurately estimating a degree of agreement requires the use of statistical logic. But if students are not sure whether a certain data set is anomalous, rather than engage in sophisticated statistical analysis, they can simply gather a lot more data, thus decreasing the need for statistical interpretation. This is another situation where the ease and speed of GCK contributes to a non-authentic science experience.

4c. Contrast: A 'crucial experiment' is rarely planned in advance. More often, after the results have been observed and interpreted, an experiment is recognized to have high discriminatory power due to the predictive contrast between an old model (which makes incorrect predictions for this experiment) and a new model (with its correct predictions). In this way a crucial experiment can be recognized in retrospect, even if it was not intentionally designed to be crucial.

4d. Previous Experiments: Students can use all experiments, both current and previous, to check for degree of agreement, and to decide whether they should reject or revise a model (especially a newly proposed model) in response to apparent anomaly. But if there are multiple anomalies, students usually try to resolve them one at a time, not all at once.

5a. Logical Structure: Conceptual evaluation criteria include systematicity with an absence of suspiciously ad hoc components, and a simplicity that avoids the use of an unusually large number of genotypes in a mechanism that is intended to causally explain a much smaller number of phenotypes. And some students are even suspicious of a model that they think is “too difficult for a high school science course.”

Students generally prefer an explanatory theory, which postulates a composition-and-operation mechanism to explain what is observed, over a descriptive theory that, even though it can make accurate predictions, does not postulate a causal mechanism. This preference for a theory that can explain, not just describe, is partly because the teacher tells the students about her own criteria for an adequate theory. More generally, this preference, by students and teachers and scientists, may be due to an innate human desire to make sense of experience by explaining why it occurs. It is also consistent with students' prior experience in which scientific theories (including the genetics models used in this course) usually postulate an explanatory cause-effect mechanism. If students use these theories as a standard for judging “what a theory should do” they will not be satisfied with a newly invented theory that is less complete because it merely predicts (by using data patterns) without supplying a mechanism.

5a-5b. Internal Consistency and External Relationships: Students begin Round 1 with an initial model (for dominance-and-meiosis) that already has been checked for the acceptability of components, and for internal and external logical consistency. During model revising, new models are invented by tinkering with part(s) of the initial model; after each revision there can be a re-checking, to ask whether the revised model is still conceptually acceptable.

The relative weighting of evaluation criteria varies from one group to another. Generally, however, empirical evaluation (especially degree of agreement) exerts more influence than conceptual evaluation. One reason to prefer empirical evaluation is that students have a limited base of knowledge about biological phenomena and theories. This limited knowledge makes it difficult for students to state confidently that a model should be rejected because it is inconsistent with known phenomena or with currently accepted scientific theories. By contrast, students can be confident about whether or not the predictions of a model agree with the data generated by GCK.

Although their overall knowledge is limited, students can use their initial models, and the principles contained in these models, as a basis for conceptual evaluation. Based on their initial model of meiosis, one commonly used conceptual principle is that each progeny must receive an equal number of alleles from each parent. In Rounds 1 and 2, this constraint is used to challenge a false model in which an individual has three alleles. But in Round 3 it can prevent students from developing the valid model that, for the trait being studied, a male has only 1 allele, always received from the mother. And in Round 2, when students don't realize the distinction between '3 alleles per individual' (not consistent with meiosis) and '3 alleles in a population' (consistent with meiosis), they may eliminate the true model from consideration, at least temporarily.

Two other commonly used conceptual criteria involve dominance. Sometimes students refuse to give up this concept without a fight, and during Rounds 2 and 3 they apply it even in contexts where the data should make it clear, based on knowledge that should have been learned in Round 1, that a particular allelic relationship is one of codominance, not dominance. Another unwarranted conceptual constraint involving dominance occurs in Round 2, when two of the three heterozygous allelic relationships (1-2, 1-3, and 2-3) are dominant-recessive, and one is codominant. In this context, despite the empirical evidence, some students find it difficult to accept the concept that each of the three relationships is independent of the others, that the relationships do not have to be either all dominant or all codominant.

When students are unwilling to revise a conceptual premise, this premise attains the status of a *protected component.* A protected component can operate at a level that is either unconscious (to prevent a perception that changing the protected component is even an option) or conscious (to decide that a component should not be changed, or to prevent the acceptance of a change that has been proposed). A component can be protected because students think it is true, or because they think that changing it will make a model less useful. For example, with a "3 alleles per individual" model, students have difficulty imagining how a Punnett Square (a familiar tool) can be used to make predictions.

A protected component can be beneficial for problem solving, or detrimental. Based on my study of students solving GCK problems, the difference between benefit and detriment seems to depend mainly on the truth of a component. Components that are true (i.e., those that correspond to what actually occurs in the simulated 'natural world' created by GCK) usually help students construct models that produce accurate predictions, and components that are false tend to hinder. The correlation between truth and predictive utility may not be 100% reliable, but it certainly is a good way to bet. In this course, based on student transcripts the correlation does seem to be 100%, although in the history of science there have been rare exceptions. Also, in real science we cannot know with certainty what is true, so the correlation that I claim cannot be proved. But it still is a good way to bet.

6a. Metaphysics: During their GCK practice the students learn that, because systems which appear identical may not actually be identical, a metaphysical assumption of consistency should be interpreted with conceptual understanding and logical sophistication. Often, patience is also required. In GCK the nature of an experimental system is not immediately apparent from observable data; instead, it must be inferred by retroductive logic, and sometimes an inference should be delayed until more data is collected. Hence the need for patience; ambiguity must be tolerated for awhile, to avoid the error of jumping to conclusions and landing on the wrong ones.

6b. Personal Factors (Psychological and Practical) and Authority: Students develop strategies for "how to do science in the classroom" based on their ideas about both science and the classroom. These ideas can be described as a context-dependent 'thought style' that operates in the classroom and in each group, that influences students' approaches to experimenting and theorizing.

During GCK problem solving, students have important interactions — both social and intellectual — with other students in their research group, and with the teacher. Students in a group can discuss not only genetics concepts, but also their options for probing, to decide which actions will be most effective for pursuing their shared goal of developing a satisfactory problem solution. Students interact with their scientific colleagues (fellow students, inside and outside their group) by competing and cooperating. Each of these modes of interaction can occur in ways that are harmful or beneficial, although cooperation is usually beneficial.

The teacher also participates in the process of learning, in her special role as expert and mentor. Even though in this course the teacher is less "authoritative" than in a conventional classroom, she still is viewed as an authority. When she asks questions or provides information or hints, students usually listen carefully because they know that Sue knows a lot about genetics and the process of scientific problem solving, she has programmed the GCK parameters so she knows “the answer” to their problem, she will be assigning the course grades, and because they know that she really does care about her students, so what she is saying is intended for their benefit. There is also authority within a group, with student's ideas being given different amounts of weight in the group's evaluations and decisions.

7a. Evaluation and Conclusion: Students can evaluate a model based on their estimates of its plausibility and/or utility, for purposes of either pursuit or acceptance. In addition to definite conclusions to retain, reject or revise, there can be an indefinite ‘delayed conclusion’ while more data is being collected and interpreted, or there may be a reversal of an earlier decision. During the process of theory evaluation, students are also actively involved in the invention and evaluation of probing actions. The remainder of "7a" examines the interactions between theory evaluation and action evaluation.

When the process of anomaly resolution is neatly summarized by focusing only on productive actions, it is simple and efficient; a new model is invented, and is evaluated as being satisfactory. But in real life, anomaly resolution is usually a messy, meandering process because students rarely invent, in their first attempt, a theory that can be quickly evaluated as being empirically and conceptually adequate. Instead, evaluation usually shows the first new model to be inadequate. Or the process of evaluation may be difficult and slow, spanning a long time in which there is no definite “yes or no” conclusion. During this period of uncertainty, students may propose other new models, including variations of their own newly invented models, and there will be an overlapping of evaluation phases for the many models that are still being considered. When there are multiple models, a group must decide which model(s) to pursue for testing and development.

According to Finkel (1993), one secret of success is to develop promising models instead of rejecting them. This is a good strategy, of course, but the obvious question arises: How does one know when a model is "promising" and when it is not? For example, if it makes incorrect predictions the status of a new model will decrease, but should the model be permanently rejected? In making this decision, sometimes conceptual criteria exert an influence. If students think a certain model makes sense conceptually, they may respond to empirical anomaly by revising the model instead of rejecting it. But if a model already has low status due to conceptual evaluation, at the first sight of anomaly it will be easier to say “well, we didn't like that model, anyway” and to quickly reject it. And students usually consider a larger set of data; if a model has made accurate predictions in previous experiments, involving a variety of systems, there may be good reason to question the validity of an apparent anomaly. People sometimes make errors, so this cause of *apparent* anomaly should always be considered. But if the anomaly is legitimate, students must decide what to do about it.

In GCK genetics, making wise decisions about whether to reject or revise, and how to revise, requires an understanding of the two foundations for prediction: a model, and a system-theory. Even if a model is correct it may produce incorrect predictions if the experimental system is not properly characterized. When students understand this dual dependence, they realize that anomaly does not mandate the automatic rejection of a model. Maybe the model just needs to be adjusted; or maybe the model is fine, but the system-theories need to be revised. In her analysis of student problem-solving behavior, Finkel (1993, p. 264) observed that successful groups "do not discard promising models but alter them slightly (for example by re-arranging the ways in which genotypes are assigned to phenotypes) until they fit the population being studied." When a model is first proposed it is often only a vague outline that needs to be developed by filling in specific details — such as genotype-phenotype mappings; and which trait-variations are dominant, recessive, or codominant; and whether certain parent-flies are heterozygous or homozygous, and with which alleles — before the model (and its application to the collection of fruitflies) can be adequately tested and perhaps revised. Or maybe the system-theories are proposed first, and from them a model is developed. Either way, students can obtain a fit with existing data by adjusting the model and/or the system-theories. And whatever strategy is used, there is usually a close connection between “knowing the genotypes” and “having an explanatory model,” with the former leading to the latter, and vice versa. For this reason, and others, when students are deciding how to test and develop a model, and whether this action should be pursued, one valuable asset is a good *working knowledge of genetics,* with fluent, accurate thinking about the connections between genotypes and phenotypes.

To solve a problem the most efficient strategy is to move on a straight-line path toward a solution. But during model revising the nature of a solution is unknown until it is found, so knowing “which direction to go” is not clear, and students usually follow a nonlinear route of goal-oriented wandering. During their exploration of multiple models, students try to balance the tension between the virtues of tenacious hard work (in developing a model that shows promise) and flexible wisdom (to stop wasting time on a model that isn't working and probably never will). If there are repeated cycles of inventing-and-testing, a useful skill is speed; if a group goes down a wrong path for awhile, this doesn't matter as much if they quickly discover that the path is not likely to be productive. Then they can return to the task of looking for another option to pursue, and repeating the cycle of exploring one potential solution after another until a satisfactory solution is found.

7b. Persuasion: Internally-oriented persuasion, inside students and within each research group, occurs frequently, for the purpose of evaluating both theories and actions. Externally-oriented persuasion, although occasionally a goal (of student or teacher) during student-teacher interactions, is not the main goal until the next activity.

B: Student Conferences

The purpose of each conference (following Rounds 1, 2 and 3) is to allow selected student groups to explain a model they have developed during their GCK work, and to present evidence supporting this model. Using a computer equipped with an overhead projector so everyone can see what is happening on the screen, this group does experiments with a new field collection of fruitflies whose inheritance is governed by the same patterns that were operating in the collections for which their model was developed. The goal of the presenters is to clearly explain why they rejected the old model(s), how they revised an old model to make a new model, and why this new model should be accepted. Usually a group's efforts to persuade will focus on experiments, supplemented by arguments that are communicated both verbally and (by drawing on the chalkboard) visually. Some preliminary planning, about what to do during the conference, can be done before the action begins. But because the presenters are working with a new field collection, they must do some improvisational “thinking on their feet” in order to apply their model in this new context. For example, they may have to determine which variations are dominant or recessive, whether certain parents have homozygous or heterozygous genotypes, and so on. More important, they are trying to quickly design experiments, based on their experience, that will provide strong empirical support for their model, in an effort to effectively persuade the listeners. Typically, a crucial experiment (with results that are anomalous for old models, but are consistent with the new model) will be effective for persuasion.

The rest of the students (and the teacher) can ask questions about the model, or they can suggest an experiment. These questions and experiments may be motivated simply by curiosity, in an effort to learn more about the model and its application; or the intent may be to challenge and falsify the model being presented. When an experiment is suggested, the presenters are expected to predict first and then do the experiment for all to see. Thus, during a conference there is an actual *pre*-diction, and therefore a public accountability that does not exist during regular GCK work when students can do an experiment first and then explain it with a prediction that is actually a *post*-diction.

During a conference the speakers try to inform and persuade, in an effort to influence the evaluations made by other students and by the teacher. In contrast to the earlier Black Box conference where the models are usually flawed in one way or another, and challenges (in the form of questions and experiments) are likely to lower the evaluative status of a model, in a GCK conference students typically do not present a model unless they are confident about its adequacy. Usually this confidence is justified and the new model is accepted by the majority of the audience. This does not mean, however, that the older models are rejected, because there can be a peaceful coexistence with the old inheritance patterns applying to some traits, and the new model applying to others. Therefore, the two main questions — "Should the old model be retained, and should the new model be accepted?" — can both be answered yes, for different situations.

3.35: Activity Group #5 — Manuscript Preparation

For this activity, which is a natural extension of the conferences that follow each round of model revising, each group chooses a model they have developed, and they write about it, by describing their model and the process (the research methods, data, and interpretations) used to develop it.

A: Manuscript Writing and Manuscript Revising

As summarized in Table 7, by looking backward and forward we can see that all preceding course activities prepare students for writing this paper, which in turn, if there is skill transfer to school and life, will help students become better scientists and communicators.

Posing? Students have more freedom of choice than in previous activities, because they can choose which of their models to write about. Thus, there is finally a "yes" in the area-posing half of the posing cell.

While students are writing, they may recognize gaps in the support for their model. These gaps can motivate new probing activities, done for the purpose of collecting or interpreting additional data. The goal of this activity is to write a paper. This requires many ‘pursuit choices’, more than when working with GCK, because there is such a wide variety of ways that a paper can be structured, in terms of both content and style, in order to achieve an objective of effective persuasion.

Although an old model is selected, this selection is not done for the purpose of obtaining a match with known data, so the "select an old theory" action is not being done. Students could try to "invent a new theory" by revising their old model during the process of writing about it, but this is not part of the assignment, and it is not expected.

Similarly, students can use old experiments but they also have the option of designing new ‘demonstration experiments’ that they think will be more impressive for persuasion. Experimental data can be taken from the group's lab notebook, or by returning to the GCK file (saved on the computer) that contains the group's previous work on a problem.

Empirical factors: An effective persuasion strategy is to show the agreement between observations and the predictions of a new model, especially if this agreement is contrasted with the incorrect predictions of the old model. Conceptual factors: When students explain what the new model is, they can show the integrity of its internal logical structure and its external consistency with currently accepted scientific theories. Cultural-personal factors: As in all persuasion, it is wise to ask “What are the most influential personal motivations for my audience?” and to plan accordingly.

Similar to the actual publishing process, students submit their papers to an editor (the teacher) to receive constructive criticism, and then revise their papers in response to this criticism. The teacher, consistent with her motivational strategy of helping students feel like real scientists, publishes student papers in a classroom journal with the official-sounding title, *The Proceedings of Monona Grove Science.*

3.4: The Second Phase of Analysis — The Structure of Instruction

Building on the foundation provided by the first phase of analysis, a second phase (in Sections 3.41-3.48) examines the functional relationships and ‘structure of instruction’ in the Monona Grove classroom.

3.41: An Introduction to the Second Phase of Analysis

The second phase of the ISM-based analysis is summarized in Table 8; this ‘supergrid’ is an overview of the entire course that shows the 17 categories of science experience for each activity analyzed in Sections 3.31-3.35, plus two additional activities: interactions with the teacher, and listening to stories. Rather than describing each of the 198 cells in detail (notice that many cells are blank), I have emphasized the features that are most relevant for understanding the structure of the classroom instruction.

The super-grid (Table 8) contains two kinds of typographical symbolism. First, in the body of the grid the most important elements are in bold-face type. Second, the activities at the top border are grouped into categories: the two **bold** activities are model revising, the *italic* activity-group generates conceptual knowledge, the two ***bold italic*** activities involve both model revision and concept generation, and the three persuading activities are outlined. The GCK practice (with no model revision) is in plain print, as are the teacher's conversations and storytelling.

Due to limitations of space with such a large number of cells in the grid, abbreviations have been used for some terms; these are listed at the bottom of the table, and are explained in the text that follows and in Section 3.48. The remainder of this section is a brief outline of the most important elements (those in bold-face type) for the purpose of introducing the topics discussed in Sections 3.42-3.47.

Table 8 (is in the "tables.pdf" file)

Most of my functional analysis will focus on "GCK model revising" because this is the central activity in the course. As shown in the "forward preparation" row of the super-grid, students prepare for their GCK model-revising work by doing similar "process" work in three preliminary activities; these preparations are discussed in Section 3.42. The same row indicates three activities in which students learn "content" that helps them prepare for GCK work. Section 3.43 discusses the main function of this preparation (to provide conceptual knowledge), plus two other instructional functions: the structuring of the Mendelian model (in Mendel's Bible) affects procedural skill, and gaps in the models (Mendelian and meiotic) produce a need to revise models during the GCK work.

The teacher takes advantage of these gaps to pose problems that are challenging yet do-able, as explained in Section 3.44, so students can have an enjoyable science experience. The importance of producing an appropriate level of difficulty is discussed in Section 3.45, as are techniques used by the teacher to adjust this level during GCK work; these improvised real-time adjustments are indicated in the "interaction with teacher" column by the "preparation as necessary" elements. Adjustments of difficulty level also occur before GCK work begins, as explained in Sections 3.42-3.44. The instructional functions served by social-intellectual interactions of students with the teacher and with other students are discussed in Section 3.46. As shown in the "listen to stories" column, by telling stories the teacher can help students increase their procedural knowledge and motivation; these functions are discussed in Section 3.47.

Finally, Section 3.48 reviews the six preceding sections (3.42-3.47) and returns to a direct examination of the functional relationships shown in the super-grid, with a detailed discussion of vertical relationships (within activities) and a summary of horizontal relationships (between activities).

3.42. Preparation by Learning Procedures

When students begin model revising with GCK, in the Black Box activity they already have done model revising in which they develop *logical strategies* for exploring an unfamiliar problem situation with observation-and-interpretation activities. With the Black Box or GCK the basic strategy is similar — recognize an anomaly (by collecting observations, making predictions, and comparing these for agreement) and resolve this anomaly (by inventing a revised model) to reach the goal-state for a solved problem. During their work on the Black Box, students also develop *psychological strategies* for coping with uncertainty when "things aren't happening as expected."

Students also gain valuable procedural experience while constructing the Mendelian and meiotic models, especially by observing two "expertly done" model developments of the type that students will be doing later with GCK. But in Table 8 these developments are labeled "model" (not worthy of capital letters) because decisions and actions are done by the whole class, with the teacher providing guidance as necessary. Success is guaranteed because it does not depend on the decisions or actions of any individual student or group, by contrast with the student-dependent invention of a MODEL during the Black Box and GCK work.

Although the most directly relevant prior experience comes from model revising activities (for the Black Box and the Initial Models), the "GCK practice without model revising" also serves an important function by helping students learn about their initial models and the GCK program.

3.43: Preparation by Learning Concepts

A. Providing Conceptual Knowledge for Model Revising

For each of the three *italicized* activities (genetics phenomena, Mendel model, meiotic model) the main instructional function is to serve as "forward preparation" by providing content-knowledge that will be used during GCK model revising. But two other functions are also served by conceptual knowledge in the Mendelian model and meiotic model.

B. Simplifying the Process of Analysis-and-Revision

The structure of conceptual knowledge can also be used to improve procedural knowledge. A key insight into the process of theory invention by analysis-and-revision comes from Wimsatt (1987): For the purpose of resolving anomalies, a model is more useful for model-revising if this model and the associated methodology (the experimental and interpretive tools that are available for analyzing the model) are structured in a way that allows the localization of anomaly to specific components of the model, which can then be revised to produce a new model that resolves the anomaly. This characteristic — a model structure that lets a user localize the anomaly and then resolve it — has been intentionally designed into Mendel's Bible, the external representation of Mendel's theory that is the initial model used by students.

Due to this preliminary structuring by the course designers, a significant part of the work of analyzing and structuring the components of Mendel's Model has been done for students, and the level of difficulty has thus been adjusted to make problems easier to solve. This structuring is especially important because it improves the effectiveness of the most commonly used tool. The primary strategy for problem solving, which the teacher encourages and students actually use, is to resolve anomalies by revising parts of Mendel's Bible, which — by serving first as a standard for defining expectations and *recognizing* anomalies, and second as a template for inventing new theories that can *resolve* these anomalies — facilitates the two main steps in model-revising problem solving. Most groups can recognize anomalies; it is in the second, more difficult step (inventing a new model to resolve the anomalies) where the structuring of Mendel's Bible is so useful for achieving problem-solving success.

In recent years, beginning in 1993, students have played a more active role in constructing Mendel's Bible in a 'guided invention' process. But this participation does not change the basic balance-adjusting principle, that before students begin model revising they have an initial model that has been neatly analyzed and structured (in ways determined by the teacher) for the purpose of making the model easier to revise.

C. Limiting What Students Know About Genetics

As discussed above, the nature of an initial model (Mendel's Bible) affects the process of solving a problem. But a problem would not even exist, except for a very important feature of the initial models — their limited information content.

One reason to limit the content is historical authenticity; the students' initial models contain what was known in the early days of classical genetics, at around 1903. But why should a course in the final decade of the twentieth century focus on science from the first decade of the century? Even though the students' knowledge of genetics is limited, with only the initial models they learn in this course they can solve problems that were frontier science in the early 1900s. But to solve problems at the frontiers of modern genetics, students would need much more knowledge. After all, problems on the modern frontiers would not still be on the frontier if typical high school students had a reasonable chance of solving them.

For instructional purposes the most important reason to limit the initial models (whether or not they are historically authentic) is to allow room for growth. Due to the limitations, students can participate in the process of developing revised models, just as the early geneticists revised their models in the years following 1903. For example, the initial Mendel's Bible, which accurately summarizes the Mendelian model of the early 1900s (including Sutton's chromosomal mechanism), involves only two alleles whose relationship is always one of dominance. These limitations allow the initial model to be revised, during Rounds 1 and 2, to include the possibilities of codominant relationships and three alleles. During Round 3 and (if there is one) Round 4, the 'meiosis' parts of Mendel's model are revised.

The key to effective guided inquiry instruction is maintaining a balance between making a problem too easy and too difficult. To produce this balance, the amount of information that is provided — both in what is taught and how it is taught — must be carefully controlled. One strategy for limiting what is learned takes advantage of a familiar difficulty in education — that students usually do not learn everything that can be learned. Therefore, some limitations on content occur when information that is available is not learned because this information is not effectively observed, interpreted, and processed into long-term working memory.

During the development of a meiotic model, for example, when the genetic determination of sex is explained, the differing lengths of X and Y chromosomes are clearly visible. Students are free to draw their own conclusions from this, but Sue does not call attention to the difference in length or its genetic implications, and students rarely remember that the Y chromosome is smaller, nor do they speculate about its significance for inheritance in males.

In another case, most students make an unwarranted generalization when they see alleles for different traits placed on different chromosomes. This is done to explain the results observed by Mendel, who did see independent assortment for all the traits he studied, so in this context the concept is correct. But it is incomplete because Sue does not call attention to the fact that for other systems (besides those studied by Mendel) the traits being observed may be determined by alleles located on the same chromosome. The placement of alleles on different chromosomes is not accompanied by the incorrect statement that this always occurs, but neither is there a warning to the contrary.

Because the course goals are for students to learn not just process skills, but also genetics content, the teacher wants students to eventually construct a unified mental model that combines an inheritance pattern (such as dominance or codominance) with the principles of meiosis. But the word "eventually" is important, because if students develop a conceptual understanding that is too complete, too soon, this conceptual knowledge will spoil their opportunity to develop procedural knowledge during the ‘process of inquiry’ experiences in the GCK model revising.

In summary: In the initial models the relationships between alleles, chromosomes, meiosis, and phenotypes are correctly specified for one type of system (the type studied by Mendel), but there is no mention of the fact that other types of systems, with other relationships, are also possible. Students are told, however, that parts of their initial models will be revised; and eventually, in the conference that follows each round of model revising, other possible relationships (such as codominance or X-linkage) are explicitly clarified in conferences when the students' newly developed models are presented and discussed.

Although the teacher can control the information provided during the course, she cannot control what students bring into the course from prior experience. Most students in the course have been exposed to genetics concepts earlier in their high school education,[23] and these concepts include all models used in the GCK problems: dominance, codominance, multiple alleles, sex linkage, and autosomal linkage. As discussed above, however, "being exposed to" is not the same as "learning and remembering." According to the teacher,

> Having worked with the genetics mini course for five years, it has become obvious that the students' memory of inheritance patterns other than simple dominance is very limited. By observing them as they construct models for these phenomena, it is easy to see that they are usually 'starting from scratch.' Occasionally a student will remember some of the names for the inheritance patterns other than simple dominance, however, if that occurs, it is usually after a model has been constructed or that they are using a name without an understanding of the model. My numerous observations to that effect come from listening to the research groups as they model revise, asking questions of individual students and groups, listening to many hours of student dialogue transcribed from audio tapes made during the process of model revision, and from the research papers produced by the groups. (Johnson, 1996, p. 32)

That students seem to be "starting from scratch" may be a sad commentary on human memory, the educational system, and/or student motivation, but in this classroom the memory loss does serve a useful purpose.

Even if students do not accidentally remember a problem solution from prior learning, they can intentionally short-circuit the process of inquiry by looking in a textbook for the revised models they are supposed to be inventing. This is a possibility, but the teacher's experience leads her to believe that this type of "inquiry cheating" occurs rarely or never, probably because students enjoy the challenge of problem solving in this course, and want to succeed by their own efforts.

3.44: Posing Problems

The question of "Who does the posing?" is examined carefully in this section, because posing is an important science experience, and because the MG classroom was designed to be a '3Ps' classroom, yet an opportunity to experience one aspect of one of the Ps seems to be missing.

---

[23]. Sue Johnson is part of a team that teaches the previous course where students could (but usually don't) learn the genetics concepts, so she does have some control over prior experience.

A. Posing is done by the Teacher

When scientists do research, they try to pose problems that are original, significant, and capable of being solved in a reasonable time with the available people and physical resources. Scientists don't work on problems that are too easy, because these usually have been figured out already. And they don't work on a problem that is too difficult, because they won't be able to make satisfactory progress.

In designing the MG course, one goal is to produce a level of difficulty — not too easy, not too difficult — that makes the students' simulated science experience similar to that of research scientists. To accomplish this function of finding problems at an appropriate level, so students will be challenged but will be motivated to work on a problem (thus turning the problem into a project) because there is hope for success, the teacher selects the problems that will be solved. The mechanism for area-posing is that when the teacher sets the parameters in GCK she creates a ‘miniature universe’ that contains only one type of natural system. Because the ‘GCK universe’ contains only one type of system, this is what the students investigate, and they soon discover the problem area that has been pre-determined by the teacher.

Thus, there is a tradeoff. If students are to have an authentic science experience with problems that are challenging yet do-able, the teacher (not students) must control the problems that will be solved. Since students do not do any area-posing, there is a "no" in the top half of the posing cells in Tables 3-6.

B. Posing is done by Students

But there is a "yes" in the bottom half of the posing cells, because students do participate in three types of posing — by defining “what needs fixing” when they recognize anomalies, by defining the constraints on what will be considered an acceptable solution, and by posing sub-problems that will contribute to solving the problem. These three types of posing — the first two are constraint-posing, while the third is area-posing at the level of sub-problems — are discussed in the following three paragraphs.

First, in order to pose a scientific problem, students must recognize one or more anomalies that indicate a weakness in the current structure of knowledge. After students have recognized what needs fixing, they can define the goal of problem solving as moving from an unsatisfactory now-state (with recognized anomalies) to a desired goal-state (with resolved anomalies and an improved state of knowledge).

Second, for the GCK problems used in this course the desired goal-state for a solved problem, as defined by the teacher, is for a group to construct a model that the students themselves judge to be empirically and conceptually adequate. Instead of the teacher judging whether a model is adequate, this responsibility is delegated to students. For example, all three students within a research group can critique a model and evaluate it, and at a conference the students in other groups can offer criticisms. The teacher does reserve the right to participate in the critical discussions of models (and of the evaluation criteria that are being used to judge the models), not as the final authority but as a colleague, in the role of a fellow research scientist. However, Sue prefers to leave the criticism-and-evaluation to students, as much as possible, only making contributions in order to fill a gap in the students' evaluations or to add commentary that will be interesting and useful.

Third, in their efforts to solve a problem the students are constantly posing sub-problems, asking questions such as “how can we revise our existing models to resolve this anomaly” or “how can we test this newly proposed model” or “how can we design an experiment that will be maximally persuasive for our conference presentation?” Consistent with the "levels of problem solving" perspective in Section 2.71_K , these questions can be viewed as ‘sub-problems’, and the search for answers requires ‘specific activities’ that function as probing actions whose purpose is to help solve a ‘problem’ that was predetermined when the teacher set the parameters in the GCK program, and was defined when students recognized anomalies in the GCK data.

In addition, area-posing occurs before the course begins, when students choose to take this elective course instead of other options that include comparative anatomy, organic chemistry, qualitative analysis, nuclear physics, dynamics of motion, electricity and magnetism, optics, and astronomy.

C. Do Students Pose Problems?

There are many legitimate answers to this question, because the response depends on how 'posing' and 'probing' are defined. An important part of this definition is the choice of a contextual perspective. The role of an action — whether it is defined as posing or probing — will depend on which problem-solving level is used to define the perspective. If the viewer sees events in the wider context of the problem, the sub-problem is part of the probing that is done to solve the problem. Or the viewer can focus on the posing and solving of the sub-problem itself.

The latter view is adopted by the formulators of the 3Ps model, in the examples they choose to illustrate problems and posing. In the section of their paper devoted to problem posing, the authors explain that during GCK work "the student's problem is: What cross to make first? Which of several traits to pursue? What hypotheses can be drawn from the initial data? Can these be confirmed or rejected? (Peterson & Jungck, 1988, p. 17)." With this view, a wide variety of questions are defined as posing.

But I think it is more accurate and useful to view these questions as potential probing actions that may be productive in the pursuit of a solution for a larger problem. Although it is possible to shift perspectives and thus to view a sub-problem as a problem, this seems like an unusual way of thinking about scientific problems and research projects. With the broad interpretation of posing suggested by Peterson and Jungck, almost any question that could ever be asked can be viewed as problem posing, and many useful distinctions disappear. But when viewed as probing — as observation and interpretation in recurring sub-goal cycles, with feedback and guidance provided by ongoing evaluation — these diverse questions fit into a flexible yet coherent 'system of activity' that can be used to describe the integrated relationships between all of the sub-problem actions that occur in pursuit of solution during a problem-solving research project.

The ideas expressed above — about what students are not told (about genetics concepts) and what they are not allowed to do (in area-posing) — are not a criticism of the course or the teacher. Instead, my analysis is intended to express an appreciation for the fine art of balance, for the difficulties involved in producing an appropriate level of difficulty in guided inquiry instruction. This analytical appreciation continues in the next section, which examines what the teacher does in the classroom.

3.45: Adjusting the Level of Difficulty

A. Why Adjustments are Important

During any type of 'guided inquiry' instruction such as that used in this genetics course, the teacher — like the writer of a good mystery story — should aim for a proper balance between the information that is provided and withheld. The goal is to give enough guidance but not too much, balanced to create an intermediate level of challenge so that students will not become bored if the task is too easy or frustrated if it is too difficult. Ideally, students will succeed, and in doing so they will feel an emotional-and-intellectual satisfaction because their success is valued as being worth something due to the obstacles they had to struggle with and overcome during the problem-solving process.

In the MG course, one of the most important functions of the 'structure of instruction' is to produce and maintain an appropriate level of difficulty for the central activity, GCK model revising. One reason, discussed earlier, is to make the students' experience similar to that of a real research scientist. Another important reason is that adjusting the difficulty level is the main factor in determining the amounts and types of thinking that students are expected to do, and are allowed to do.

B. When to adjust? Before or During Problem Solving

The level of difficulty can be adjusted at two times: before a problem begins, and while students are solving the problem.

Before the GCK model revising activity begins, the difficulty level can be adjusted in four ways: by giving students prior experience in solving similar problems, by supplying a model that is structured so it facilitates analysis-and-revision, by limiting the information that is provided, and by selecting the phenomena that will be studied (which determines the problems that will be solved). These preliminary adjustments were discussed in Sections 3.42-3.44.

Adjustments also occur after a problem begins, with the teacher playing a key role in adjusting the level of difficulty by providing assistance while students are solving a problem. It is important to recognize that the teacher's improvised coaching is intentional, not due to a need to "fix what should have been done better." Coaching is an integral part of the course design and the classroom instruction. Initially, the problem difficulty is set a little higher than most students can cope with; then Sue provides customized help, if and when it is needed, during the process of problem solving. These interactions also provide an opportunity for learning that can be both conceptual and procedural, intellectual and emotional.

In her interactions with students, the teacher plays several roles: she serves as a source of information about problem solving and biology, an adjuster of problem difficulty, a source of emotional support, and a facilitator of learning. These five roles are described in the next five subsections.

C. The Teacher as a Source of Procedural Knowledge

During her introduction to model-revising activities, the teacher provides general guidance by explaining to students what they will be doing, how they will do it, and why. Following this, during the process of problem solving she provides personally customized guidance as she circulates around the room, visiting each research group to see how they are doing, to answer questions and ask questions, to offer intellectual advice and emotional support.

One type of assistance is procedural; if students are unsure about "what to do next," the teacher can suggest ideas. Usually, to let students do as much as possible by themselves, this advice is kept to a minimum. The hope is that the procedural knowledge gained from experience in the previous activities will be sufficient preparation. Or, to avoid giving away crucial information, the advice is helpful yet general; for example, the most common suggestion, especially for model invention, is to "use your Mendel Bible" as a reminder about the components that can be modified.

D. The Teacher as a Source of Conceptual Knowledge

Because Mendel's Bible is an organized summary of the initial models, if a group is not making problem-solving progress because students do not understand their genetics models, Sue may offer the general advice to “use your Mendel Bible.” Or students may have specific questions, which the teacher answers carefully, in accordance with her goal of maintaining an appropriate level of difficulty.

Another useful function of the teacher is necessary because students have a limited base of knowledge about biological phenomena and theories, so they often cannot make appropriate conceptual evaluations — for example, by asking whether a newly proposed model is consistent with currently accepted theories. By supplementing the students' limited knowledge with information that is provided on a “need to know” basis, the teacher adjusts the students' situation so it more closely resembles that of a typical genetics researcher, who would have a broad foundation of knowledge about biology. In providing this background knowledge, however, the teacher is always concerned with the need to maintain a balance by giving information without giving away the answer.

E. The Teacher as an Adjuster of Problem Difficulty

Throughout the course, one of the teacher's main goals is to maintain a pedagogically effective level of problem difficulty. One way to do this is to select the problems that students will solve. Another way is to make fine-tuning adjustments while a class is in progress. These improvised adjustments, customized to meet the needs of individual groups, occur in real-time interactions as the teacher walks around the room and talks with individual research groups, watching and listening, and giving whatever customized assistance she feels will be beneficial to the students, both intellectually and emotionally.

Students want to have fun during the process of solving, and to feel satisfaction when they find a solution. But to feel the emotions of fun and satisfaction requires a challenging problem, with students doing as much as possible for themselves. The teacher, guided by these principles and her awareness of each group's ‘state of mind’ and their ‘state of progress’ toward a solution, practices the fine art of providing the types and amounts of guidance that will be “just right” for a particular situation. The immediate goal is to help students get moving along a path to their own solution, but without giving so much help that it spoils the opportunity for fun and satisfaction.

Students agree with this balance of guidance. They do ask for help, but they want the type of assistance that maintains a suitable level of problem-solving difficulty. They want to earn the answer, not have it given to them.

F. The Teacher as a Source of Emotional Support

The previous subsection emphasized the importance of avoiding the error of "too much help." But, of course, the reverse error is "not enough help." Students want to succeed, and consistent failure can have a devastating effect on confidence and motivation. After too much failure, students may give up hope and give up trying — instead of pursuing a solution to the GCK problem that patiently awaits their attention on the computer screen, they spend most of their time off-task, talking about everything except genetics.

To help students avoid this emotional response, the teacher can try to provide some assistance (but not too much) so a group will experience success and feel good about it. But Sue also plays an important role as a source of emotional support. A common function of the teacher — to console and motivate the students in a group that is temporarily experiencing failure — is illustrated in the following exhortation by the teacher:

> 'How many times does a scientist go with an idea in mind and, bingo, that's what it turns out to be? It's probably very unusual, actually. OK, what were some of the other things you toyed with?' (Lemberger, 1995, p. 271)

Sue comforts the students, and then nudges them back into action.

One reason to provide support is to avoid a decrease of motivation during classroom activities. But there is a more important reason for encouraging students to persevere. In the long run, one of the most valuable lessons a student can learn is to keep working despite periods when the effort doesn't seem to be paying off. An effective way to learn this skill is by realistic practice, by working through the tough times and experiencing the success that results from this investment of effort. Ideally, every student will experience a combination of hard work and satisfaction. In reality, this may not happen. Some students will not be challenged enough, while the problems are too tough for others. But by trying to maintain a balance, through a combination of preliminary planning and adjustments during problem solving, personally customized for each group, the teacher hopes the classroom experiences will provide "the greatest benefit for the greatest number of students."

3.46: Helping Students Learn from Their Experience

During all classroom activities, each student is engaged in frequent social and intellectual interactions with other students and with the teacher. For example, while a group is inventing and evaluating potential probing actions, the group members usually talk with each other. These interactions are themselves 'science experiences', but this section will focus on the functions of these interactions in helping students learn from their other experiences in the course.

A. The Teacher as a Facilitator of Learning

During student-teacher interactions, questions can be asked and answered, along with discussions about the problem and how the group is trying to solve it, and whether this is producing satisfactory progress. The teacher can serve as a source of "instant preparation," providing knowledge (about concepts and procedures) as needed during problem solving. These interactions can help students learn from their experience, to improve their procedural and conceptual knowledge.

Finkel (1993, 1996) describes the ways in which three types of knowledge — of genetics, of model revising, and of one's own actions — contribute to problem solving, and are learned during the process of problem solving. A key part of the learning process is students' interactions with their scientific colleagues, which include other students and the teacher. The following summary, based on Finkel (1993), describes these interactions and what the teacher teaches:

1. Student/student interactions are perceived, by students, as a collaborative effort among equals. But when students interact with the teacher, who has greater knowledge (of genetics and problem solving) and a position of power in the classroom, there is a difference in perceived status. Student/teacher interactions can be negotiative or instructional, motivated and oriented by the goals of students and teachers, respectively.
2. The teacher helps students learn by serving as a model of expert problem-solving behavior, by encouraging students to recognize anomalies and to clarify and articulate their ideas, and by suggesting possible pursuit-actions or strategies for deciding what to do next.

B. Learning by Metacognitive Reflection

The teacher, in a variety of ways, encourages metacognitive activity. Finkel (1996, p. 362) describes a general instructional strategy:

> The teacher supported the development of students' metacognitive knowledge by encouraging them to reflect on their work, through asking students to explain and clarify their ideas. ... These requests, along with the expectation that students collaborate with one another within their research groups, supported the development of metacognitive skills. (Finkel, 1996, p. 362)

What are the functions of metacognition, in helping students solve problems and learn science?

> Metacognitive knowledge consists of students' understanding of what they are doing as they model-revise. ... It allows students to keep track of what they have done and what they intend to do next. ... [And it] helps them make connections between knowledge of genetics and knowledge of model revising, and thus helps students develop stronger conceptual understandings of genetics and model revising. (Finkel, 1996, p. 362)

Metacognitive reflection can be done by individual students or as a group.

C. Learning from Other Students

Important social and intellectual interactions occur between students, especially within each research group. While students in a group are solving a problem, an important, practical topic is a metacognitively oriented discussion of procedural strategies:

> Students use metacognitive knowledge as a way to discuss their work with other group members. Frequently conversations among peers revolve around a review of data collected...or around a discussion of strategies which have been used or might be used to develop and/or test new models. (Finkel, 1993, p. 293)

The learning experiences provided by these discussions are a consequence of an important decision in designing the MG course — i.e., to have students work in groups. There are many reasons to use groups. Some advantages are practical; with groups of 3 students, only 8 computers are needed, not 24; and the teacher can spend more time per group when her time is split 8 ways, instead of 24. The students usually enjoy working together, and they always learn from each other. Learning can occur during any of the interactions — as students persuade (regarding theories or procedures), question or explain, challenge or agree, propose new ideas, and delegate responsibilities — that are a natural part of a group's problem solving efforts (Finkel, 1993). Students can also learn the benefits of cooperation; sometimes a problem that would be too difficult for any of the students individually can be solved by the group when the individual resources are synergistically combined.

Cooperation and competition can be either beneficial or harmful. Usually cooperation is beneficial. But if there is an excessive desire for teamwork, and if this leads to an atmosphere that discourages a mutual criticism of ideas, the result may be an ineffective, uncritical ‘group think’ conformity. And although competition can lead to prideful attitudes that hinder the cohesive teamwork of a group's problem-solving efforts, and that decrease the objectivity of evaluation, competition can be beneficial by providing motivation for a group (competing with other groups) or for individuals (competing with other individuals).

Ideally, the dynamics of a group will produce a feeling of “group ownership” for ideas, where instead of an adversarial attitude of “my ideas versus your ideas” there is a cooperative effort toward doing whatever is needed to solve the problem, no matter where the ideas come from. Usually, whether or not a group is working well together, there will be differing amounts of ‘authority’, with one student's ideas being given more weight than the ideas of another. In terms of immediate problem-solving effectiveness, this is a benefit if the leader is leading the group in a productive direction, or a detriment if the direction is not productive. Educationally, strong leadership is beneficial if it helps others learn, but is detrimental to the other students if they adopt a “Why bother?” attitude and let the leader do the thinking. In this case a non-leader who does not try to contribute, or is not able to contribute even though motivated to do so, will miss valuable opportunities for learning, compared with working alone.

Usually, however, students work well in groups, despite an occasional mismatch of abilities or thinking styles, or even personal conflicts. Although a mismatch or conflict can be unpleasant, this also provide an opportunity to learn the valuable skill of working productively in spite of initially adverse circumstances. All things considered, it seems that working in groups is beneficial for most students, in most situations.

3.47: Stories about Science and Scientists

In the MG course, students learn about scientific inquiry in two main ways. The primary method is through their personal experience of “doing science” by solving GCK problems. A secondary method of teaching inquiry uses the history of science — true stories about the adventures of research scientists as they pursue their goal of advancing the frontiers of knowledge. These stories, involving scientists both past and present, provide the opportunity for second-hand, vicarious experience that complements the students' first-hand, personal experience with GCK. When students solve GCK problems, they are actively involved in the process of inquiry, but this experience is limited in scope. If GCK activity is supplemented by the history of how scientists actually solved similar problems, the students' range of experience is significantly widened to include many factors that are an important part of real-life science but are difficult to simulate in classroom science. In balanced combination, a student's total experience — personal first-hand and vicarious second-hand — will complement each other to produce a stimulating, enjoyable blend that more closely approximates the experience of a professional scientist actively involved in research.

Listening to stories can be considered either a sub-activity that occurs in the context of the major course activities, or an activity all by itself. In either case, the usual function is to support what students are learning during their first-hand experience.

This section elaborates themes introduced in Sections $3.24_{A\text{-}B}$ by examining the instructional functions of stories told by the teacher. These stories can be shared with the whole class at once during discussions or conferences, or with one group in a conversation during problem solving. In this section, my intended meaning for ‘story’ is the first definition from the Random House Dictionary (1980), "a written or spoken account of something that has happened," rather than the second definition, "a fictitious tale, shorter than a novel." In other words, I mean ‘story’ as in ‘history’.

A. Stories about Science: Strategies for Problem Solving

Sometimes the teacher uses a history of science episode to serve several pedagogical functions. Usually a story produces second-hand experiences, but these are combined with first-hand experience during the visit by Mendel. This activity provides fascinating insight into science as it was performed long ago in an exotic culture; Mendel's research exemplifies the creative process of carefully posing a problem, planning experiments, and interpreting the results using statistics (which required creative innovation because mathematical analysis was rarely used in biology in the 1800s). In addition, the activity gives students a first-hand opportunity to examine Mendel's peas and classify them according to phenotypes, and construct a model that adequately explains these observations.

Later in the course, while talking with a group during their work in Round 2 of model revising, the teacher uses Mendel to illustrate the importance of designing experiments that increase the probability of effective retroductive invention will be effective.

> You might not want to be like Darwin, cause Darwin looked at so many things in so long a time that he got all confused and was never able to come up with a model [for genetics], whereas Mendel really narrowed his research down to these little peas and their seven characteristics. Was able to pull out a model much more easily. So what you might want to do is maybe not try to figure out every single cross at one time, you might want to say the first thing you noticed was four variations. How could you...fiddle around with these ideas up here...to get more than three variations? (Lemberger, 1995, p. 217)

And if the question is what to do after experiments have been done, Mendel's actions can be used to illustrate the strategy of coping with multiple empirical constraints by trying to satisfy them one at a time. In a similar way, another story is used for another strategy:

> Again, as often happens in science, you don't start from scratch. Nobel laureates Thomas Hunt Morgan, Brige — you know, a lot of the other scientists around that time, Bridges,...there's tons of them that were in the early part of this century. They weren't starting from scratch. They were using Mendel's ideas, and then just revising them to fit what they saw that varied from what Mendel had seen. And you're going to do the same thing. (Lemberger, 1995, p. 85)

These two examples illustrate a psychologically effective instructional technique. By attributing a certain problem-solving strategy to a great scientist, who models the effective use of this strategy in actual research science, the teacher can use ‘authority’ to increase the status of the suggested procedural knowledge. With the story, it is not just Sue who suggests the strategy, it is Sue backed by Mendel or Morgan. This technique is analogous to the use of authority to support a theory, but in this case the authority is used to support procedural knowledge rather than theoretical knowledge.

B. Stories about Science: Having Fun as a Scientist

An important function of the teacher is to provide emotional support, to exhort students to persevere when they are vulnerable to discouragement because things are not going well.

> What are the pieces? Just like remember when Watson and Crick, I mean finally they were not able to get their final answer until just wham that piece of information about the different forms of that base came, you know. It may be that you're just waiting for this, you know, this little piece of information is going to get thrown into the pot here and it's going to make all the difference in your thinking. I mean Watson just sat there and bingo there it was. And...that same thing often happens. (Lemberger, 1995, p. 172)

This story performs several functions. It encourages students to persevere, it teaches a lesson about retroduction using multiple data sources (that may or may not yet include the information that "makes all the difference"), and it is a reminder of the great joy that Watson and Crick felt when they finally put all the pieces together.

Many of the teacher's stories highlight the positive aspects of real-world science. This helps her achieve one of the main goals of the course, to show students that science can be fun. In creating a classroom atmosphere that is enjoyable and motivating, many techniques are used; some of these — such as naming student groups after famous scientists (both male and female) who can potentially serve as role models, receiving name badges, participating in a Nobel Prize ceremony, publishing a paper, and collaborating with Mendel to construct the foundation theory of classical genetics — were described earlier. In addition, motivation and understanding can come from stories about science and scientists. These techniques and stories may inspire students who will become scientists; and those who do not choose a career in science may develop a deeper understanding of science and scientists.

Some students will be inspired by the possibility of contributing to medical advances based on genetics knowledge, so students read (or watch) *Alex,* the emotionally appealing story of a young girl and her family and a crippling genetic disease. Some students will be impressed by real-life applications of science, outside the laboratory, that bring benefits to humanity. Others are impressed by money. The story of Hector Deluca illustrates the financial rewards sometimes associated with advances in science and technology; in this case, Deluca and the University of Wisconsin have earned many millions of dollars from a patent for synthesizing Vitamin D. The lure of fame can also be a powerful motivator. This is illustrated by James Watson who — motivated by a desire for fame and fortune, a Nobel Prize, and a place in history — decided with a single-minded determination that he would avoid working on lesser problems so he could focus on DNA, an area "where the action was" in the early 1950s.

Sometimes rebellion is fun, especially for teenagers. When Howard Temin proposed a 'reverse transcriptase' enzyme that translates RNA into DNA, he was violating the Central Dogma of molecular biology, the principle that the flow of information occurs in only one direction, DNA → RNA → proteins, never vice versa. Partly due to this dogma, initially his proposal was vigorously resisted, but eventually it was accepted when the strong conceptual constraint was overcome by empirical evidence.

Most people enjoy a good story, and one of the more exciting stories in the history of science is revealed and interpreted in *The Double Helix* (Watson, 1968). The dramatic adventures of Watson and Crick show the intensely human aspects of science: the complex personal motivations and social interactions that occur during cooperation and competition, the frustrating but highly motivating effects of wanting-to-know yet not-knowing, the emotional roller coaster of excitement (after formulating a promising new theory) and disappointment (when evaluation shows it to be inadequate), and finally — after a great deal of time and worry, work and luck — the immense joy of successfully solving the puzzle of DNA.

The great "DNA chase" had high stakes and high drama. There was cooperation (especially between Watson and Crick) plus fierce competition, and even some espionage! Much attention was paid to priority of discovery, because this was the key to winning The Prize. This attitude of inter-group competition contrasts sharply with the feeling of intra-group cooperative teamwork inside the Morgan lab during the 1910s when many essential components of classical genetics were being developed.

> The group worked cohesively, but each individual had his own particular experiments and interests. ... The individual abilities complemented each other, providing a cohesiveness in those early years that allowed the work to progress with astounding rapidity. ... [And, quoting Sturtevant, a scientist who worked in the lab,] 'The group worked as a unit. Each carried on his own experiments, but each knew exactly what the others were doing, and each new result was freely discussed. There was little attention paid to priority or to the source of new ideas or new interpretations. What mattered was to get ahead with the work. There was much to be done; there were many new ideas to be tested, and many new experimental techniques to be developed. There can have been few times and places in scientific laboratories with such an atmosphere of excitement and with such a record of sustained enthusiasm.' (Allen, 1978, pp. 62-64)

And both of these contexts (for Watson and Morgan) contrast with the working environment of Mendel, who developed his theory alone, with minimal cooperation or competition, isolated from other scientists, while living in a monastery.

Most stories used in this classroom come from the history of genetics, both early and modern. The students' first-hand experience is with classical genetics, so it makes sense to use second-hand experiences that involve similar subject matter. If a story is from the early classical period, students have struggled with some of the same problems that puzzled these scientists, so the history will be more interesting — like a student finding a new friend who, just like her, also worked as a lifeguard at a lake last summer — and the shared experiences will make it easier to develop new insights about science. And if a story is about modern genetics, students can see some of the amazing feats accomplished by geneticists since the early days of their field.

Because history is complex and multi-faceted, the same story can be used to illustrate several ideas. In some cases, such as Mendel or Watson, it is easy to find a number of different lessons to learn, in addition to just enjoying the story. Other examples of multiple lessons include: Hector Deluca as an example of financial rewards and also to illustrate (by negative example) the importance of keeping detailed laboratory notes, which in this case would have helped the University of Wisconsin defend against a lawsuit challenging their rights to the Vitamin D patent; medical applications of genetics can raise questions about the potential rewards and dangers of genetics technologies, about ethics and our decisions (as individuals and as a society) in posing scientific problems and determining public policies. These questions and many others can be asked, by teacher or student, for any story. When used creatively in the classroom, stories — due to their connections with shared human experience, and their inherent richness — can inspire many stimulating ideas and discussions.

3.48: Functional Relationships in the Instruction

A. Functional Relationships Within Activities

In Table 8, scanning each activity-column vertically shows the most important experiences that occur in this activity. In the following discussion, which begins with the central activity in the course, an activity will often be described by comparing it with other activities. Therefore, even though in this subsection the focus is on 'vertical' relationships (within activities), these characterizations will involve a consideration of 'horizontal' relationships (between activities). One horizontal relationship — a grouping into "types of activity," described in Section 3.41 and symbolized in the font styles (bold for model revising, italic for content, outline for persuasion) — will be assumed without further comment.

GCK Model Revising. Some experiences, shown in bold print, are considered (by me, for purposes of this analysis) to be more important than others.[24] An important general experience is to face an anomalous situation and learn how to cope with it. The general strategy for coping is to pursue a solution by inventing, evaluating and executing probing actions, or (focusing on the problem-solving functions of these actions) by recognizing anomalies and resolving them. A specific strategy is to resolve anomalies by revising previously existing models, to retroductively invent a model that can explain the known data. An essential step in retroduction is to make predictions based on a newly invented model.

The title of this section is "functional relationships," yet nothing in the structure of the grid

---

[24]. Of course, judgments about importance are subjective, and depend on the evaluation criteria being used. For example, experimental design is common in this course, although it rarely is done in conventional science education. Therefore, when judged by a criterion of whether it is "a rare opportunity for a science experience," this is a valuable aspect of the MG instruction. But if a criterion asks whether students "do thought-experiments as a screening process for potential physical-experiments," the GCK work is much less valuable. By an "importance in problem solving" criterion, data collection is necessary so it is important. But since most groups can do this action satisfactorily, with success depending mostly on data interpretation rather than data collection, skill in experimenting is not very important when evaluated using a criterion of "determining success in problem solving." Obviously, a perception of importance depends on the criteria by which an experience is judged to be important. To make things more complicated, importance varies with context. During Rounds 1 and 2, for example, collecting data is necessary and commonly used, but is not a determiner of success. By contrast, for problems involving X-linked traits in Round 3, another aspect of experimenting — the ability to "mentally observe" a previously unimportant characteristic (that phenotype ratios differ by sex) — is crucial for success, even though in previous rounds this skill was never required, and was probably never used. But since trying to explain the criteria by which each element is considered important would detract from the main goals of my analysis, usually an explicit explanation will not be provided.

(except the grouping of experiences into 7 categories) indicates any kind of relationship within an activity. For this, the analysis depends on outside knowledge, on the integrated relationships in ISM. Because the 17 science experiences are logically organized according to *'function in science'* relationships in ISM, it is easier to understand the functional relationships between actions that occur in each classroom activity, to see how these actions fulfill related functions in a coherent, goal-oriented effort to solve a scientific problem. This 'transitive logic', which uses ISM to bridge the gap between students and scientists, assumes a similarity (or analogy) between the functions of student actions in the classroom and scientist actions in a laboratory. Based on studies of the classroom by myself and by other researchers, this assumption about function seems to be justified.

Black Box Model Revising. This activity is analogous to GCK model revising in the most important aspects of experience — coping with anomaly by pursuing a solution in which the anomaly has been resolved through probing actions that include the inventing and testing of new models. These similarities let the Black Box serve as a preparation for GCK.

One important difference is that, compared with the rigid limitations on GCK experiments (all students can do is select which parents to cross), there is more room for creativity in designing experiments with a Black Box.

Other differences also exist, but these do not seem functionally important. For example, with a Black Box the students do physical experiments and make observations with their senses, while with GCK they mentally observe a data summary on the computer screen. For purposes of problem solving, however, the functions are similar because in both cases there is a gathering of data that can be interpreted. And because physical-sensory experiments (as with the Black Box) are common in conventional science labs, with GCK there is no loss of "an opportunity for a rare experience." And during the course there are opportunities for sensory observation, with the Black Box and with cookies, human variations, Mendel's peas, and real fruitflies.

The Initial Models. The process of developing Mendel's model and a meiotic model is similar to GCK model revising. The major difference is self-reliance. With the initial models the whole process is guided by the teacher, and a conclusion is reached by the class as a whole. But with GCK each group works independently; the teacher provides guidance as needed, but the pursuit of a solution is the students' responsibility, and success (or failure) depends on student actions, in contrast with the initial models where success is guaranteed for all.

GCK without model revising. This is like GCK model revising, without the revising. During this activity, one new and stimulating intellectual experience is the realization that 'consistency' and 'reproducible results' must be interpreted carefully, with sophisticated logic and patience, because experimental systems that are apparently identical (but not really identical at a level that cannot be observed) can produce different results.

The Persuading Activities. These activities (Black Box conference, GCK conferences, and GCK Manuscript Preparation) differ from other activities in several ways. First, the problem to be solved (and the goal to achieve) is effective externally-oriented persuasion, not the development of a scientifically adequate model. This difference is shown in the "probing" row of Table 8. The orientation of persuasion also differs, as shown in the "persuasion" and "conclusion" rows. With the Black Box or GCK, there is internal persuasion (occurring within a group) to reach a conclusion (by the group) about a model or about pursuit-actions; basically, a group persuades itself. With the initial models, the overall result is that the teacher tries to persuade the students (as a whole class, all at once). The persuading activities are similar, except that now students persuade other students, and there is more of an "us persuading them" feeling, compared with the initial models persuasion which has an "us persuading ourselves" feeling due to the students' active involvement.

But similar criteria are used to evaluate a successful scientific solution and to plan a successful persuasion, because the goal for both activities is to decide whether a model is scientifically plausible and/or useful. In each activity the main criterion used by students is empirical adequacy, by checking for agreement between predictions and observations. This is shown in the "agreement" and "contrast" rows, which emphasize that for students the main criterion (or argument) is empirical adequacy. Due to this emphasis, an effective strategy for persuasion is to design an impressive 'demonstration' experiment that clearly shows, in a situation with high predictive contrast, the difference between an old model (with anomaly) and a new model (with agreement). In an effort to select or invent suitable experiments, a group can draw on its experience in the model-revising activity that precedes (and is a preparation for) each persuasion activity.

As indicated by the "yes, yes" in the posing cell for Manuscript Preparation, in the persuading activities there is quite a bit of freedom in selecting content and style during the fine art of effective persuasion. The choice of pursuit actions (done with the goal of finding a solution for a GCK problem) is also open-ended, with many options for problem-solving strategies and tactics. Compared with the relatively open-ended nature of persuasion strategies and pursuit actions, however, only a narrow range of models can be evaluated (by using the typical scientific criteria) as acceptable solutions for the GCK problems.

B. Functional Relationships Between Activities

Most of the ideas in the following discussion have already been discussed in Sections 3.42-3.47; these ideas will now be reviewed and summarized, from a broad overview perspective. As in the “vertical relationships” section above, my discussion begins with (and will focus on) the central activity.

In Table 8, scanning a row horizontally can reveal functional relationships, such as ‘repeated experiences’, between activities. Several important functional relationships are shown in the "forward preparation" row. To help students learn the procedural and/or conceptual knowledge they will need for GCK Model Revising, preparation occurs in the Black Box Model Revising, and also in the Mendel Model and Meiotic Model activities. Each of these preliminary activities has its own intrinsic value, but the main function in the context of the MG course is to help students prepare for their GCK work.

Due to a similarity between the types of problems that are posed for students, by the teacher, in the Black Box model revising and GCK model revising, in these two activities there is a repeated experience for a whole set of functionally analogous actions — i.e., actions that perform analogous functions in each activity. In fact, the main reason for including the Black Box activity in the course is that it provides a “set of repeated experiences” that lets students practice similar procedural skills in their study of two very different systems — a physical black box, and computer-simulated fruitfly genetics. It is especially valuable to practice this particular set of skills — the art of inventing, evaluating, and executing problem-solving actions — because even though these skills are an essential part of scientific research, they are rarely experienced in conventional instruction. In my opinion, these ‘probing actions’ are the most important general science experience, so there is an emphatic "YES!" in the probing cell for each activity.

The Black Box and GCK are similar in many ways, but one difference between them is essential for the structure of instruction in the MG course. The Black Box does not require any specialized background knowledge, so students can do this on the first day, with only a brief introduction in which the teacher explains that “what you will do in this activity is similar to what scientists do when they investigate mysterious phenomena.” By contrast, the GCK work requires preparation by learning the initial genetics models.

There is also functional analogy between the procedures used in the Initial Models and GCK activities. The main difference is that the Initial Models work is more carefully guided, and it functions as an example of “expert problem solving” for the same type of task — developing an explanatory model for initially puzzling phenomena — that students will do later in their GCK model revising.

The activity of GCK Practice (without model revising) provides procedural experience with GCK, and also helps students learn content. Because this is the only place where the absence of an experience is crucial, there is a "—" in the model invention cell for the GCK model-using activity. This absence of invention is an important aspect of the course design, because it lets students focus on learning the GCK program and developing a deeper understanding of their initial genetics models, without being distracted by the surprises that occur in their later GCK work.

In the "forward preparation" row, four activities provide "content" as preparation for GCK work. The main function of the Initial Models is to provide conceptual knowledge, but because this content will be used during problem solving there is a close connection between content and process. For example, the structuring of Mendel's Bible (content) plays a crucial role in facilitating model invention by analysis-and-revision (process).

There are intentional “content gaps” in the initial genetics models. These gaps, which are carefully planned by the course designers, allow the posing of GCK problems that necessitate a revising of the initial models to fill the gaps. Because designing an educationally useful problem is so difficult and important, the teacher decides which problems will be solved. In doing this there is a tradeoff; in order to provide one set of science experiences (with an appropriate level of difficulty, similar to what might be found in a research lab, with problems that are challenging yet capable of being solved), it is necessary to deny students the opportunity for another experience — the action of area-posing, of deciding which systems will be studied.

After a problem begins, the teacher can adjust the difficulty level of a problem by supplying information (regarding either conceptual or procedural knowledge) to each group, improvised and customized as the situation seems to warrant. This function is shown in the "interaction with teacher" column.

Since “difficulty” depends on the problem and the solver, a teacher can adjust the difficulty level of a problem in two ways: by changing the problem's intrinsic difficulty, or by changing the skills, experience, or knowledge of the solvers. Both adjustment methods are used in the classroom; adjusting the intrinsic difficulty is done by setting the GCK parameters, and the solver's skill is adjusted in many ways: by providing prior procedural experience (by doing analogous free-form model revising for a black box, by participating in the guided development of models for Mendelian genetics and for meiosis, and by working with GCK and Mendel's Model to solve problems that don't require model revision), and controlling what is known (to produce gaps in content), providing a concept-organizing structure in Mendel's Bible, and “coaching” students while they are trying to solve the problem.

Of course, the teacher is more than just an adjuster of difficulty. During her individualized coaching, and in whole-class settings, she also helps students learn conceptual and procedural knowledge. And she provides emotional support (by empathizing, exhorting, and consoling) to improve student motivation; this is indicated in the "interaction with teacher" column in Table 8. Earlier in the "personal" row, "bonding" indicates an important function of the Genetics Phenomena activities (cookies, variations, pedigrees) — to promote social bonding between students, to produce a more enjoyable atmosphere in the classroom. And at the end of the row there is another "motivation" because Sue's science stories are intended to show students that doing science can be fun, and also (as indicated further up in the "listen to stories" column) to help them learn procedural knowledge.

Finally, the analysis concludes with a reminder about the two main functions of the MG course — to provide opportunities for experience, and to help students learn from their experience. Most of the analysis has been about promoting experience, but an important function of the instruction is also to help students learn from their experience. Three instructional techniques that contribute to this function are: increasing the social-intellectual interactions between students (by having them work in groups, for example), making frequent student-teacher interactions, and encouraging metacognitive reflection.

3.5: Suggestions for Improving the Course

This section contains suggestions for improving the MG course, made by others (in 3.51) and by myself (in 3.52).

3.51: Suggestions by Others

From the beginning of their designing efforts in 1988 the course developers (Susan Johnson, James Stewart, and Robert Hafner) were motivated by a desire to change the course in any way that would improve it. This flexible attitude continued after the basic instructional structure was established, and the course has continually evolved in response to ideas by the developers and from other sources. For example, during her research in 1992 Finkel suggested spending more time to develop the meiotic model in more depth, and the following year this was done.

Changes in the early years of the course development are too numerous to mention here. Instead, I will describe (and comment on) some recent changes suggested by the teacher, inspired by her research on the strategies used by students during model-revising problem solving. The main objective of these suggestions is "to modify the course in order to generally increase student success at model revision and specifically to increase the repertoire of strategies that the students use as they revise the Mendelian model. (Johnson, 1996, p. 222)"   One way to do this would be to use class conferences to discuss not just new models, but also the strategies that have been used (or could be used) for revising and testing models. Included among the ideas to be discussed would be those of Darden (1991) who, based on her study of scientists who originally developed classical genetics, suggests a variety of useful strategies for producing new ideas, assessing tentative models, and resolving anomalies.

The suggested changes would also aim at improving students' knowledge of genetics. Due to the close relationship between content knowledge and procedural knowledge, usually this would also produce improvements in procedural skill. To help students gain a more complete mastery of concepts such as genotype-phenotype relationships, the teacher could more explicitly emphasize some aspects of existing models. And students, in order to better understand their newly developing models, could keep a summary chart for the characteristics of each newly proposed model, including its conceptual and empirical strengths and weaknesses.[25] Then, following each conference, all students would return to their computers to practice applying a "newly accepted model" in order to solidly establish their understanding of it, to make it a part of their problem-solving repertoire.[26]

I think these changes would be educationally useful, but only if they were done carefully, with caution, to avoid making problems too easy. For example, Johnson suggests that in discussing the parts of Mendel's Bible that can be revised, one possible benefit might be "to help alleviate the tenaciousness, seen in this study, with which they hold on to the dominance portion of the Mendel model. (Johnson, 1996, pp. 223-224)" But I question whether this would be educationally beneficial, in the long run. There can be some value in letting students struggle with their 'mental block' toward protecting dominance from revision in Round 1. If students can recognize that they already have been a victim of their own rigid thinking, they are more likely to appreciate the advice to "examine each constraint, and ask whether it is really necessary," instead of saying to themselves, "I would never do anything like that." On the other hand, discussing model components provides an excellent opportunity to learn genetics concepts. All things considered, I think this is a good idea if the discussion is done well, with good timing (maybe by delaying it until

---

25. Making these charts would make the students' work more similar to actual research, where scientists usually think carefully about their theories.

26. This opportunity for application, which serves an instructional function similar to the GCK practice problems that precede the GCK model revising, might help to break a 'vicious cycle' in which an unsuccessful group fails to master a tool (the new model) that is needed to succeed in the next round, thus making failure in this next round more likely, and so on. A similar purpose is served by a current course practice, adopted following a suggestion by Finkel (1993), to develop a 'bible' for each new model because this helps to give the new model a higher status, so students will tend to learn it more thoroughly, and use it more often in future rounds.

after Round 1), without lowering the level of problem difficulty. And the teacher certainly has the experience and teaching skill that are needed to do it well.

3.52: My Suggestions for Improvement

The suggestions that follow are offered with humility, due to my respect for the overall quality of the course design, and for Sue's excellent teaching. Many other people also think highly of the course. Students are consistently enthusiastic. And due partly to her work in developing and teaching this course, Sue Johnson was named the 1990 "Wisconsin Biology Teacher of the Year" by the National Association of Biology Teachers. The course has been researched extensively (in five previous doctoral dissertations, plus my own), is the main topic for six papers and two book chapters, and is described as a ‘model classroom’ in the National Standards for Science Education (1996).

In my opinion, every student would benefit greatly from taking this type of course at least once, because it provides an opportunity for experience that is productive yet is unfortunately rare in education. This experience is valuable for students in five ways. First, they learn skills in effect-to-cause problem solving and, more generally, in coping with an unstructured problem situation where “what to do next” is often not clear. Second, they learn about the nature of research science. Third, the experience is usually highly motivational for students, whether or not they plan a career in science. Fourth, students gain a deep, structured, functional knowledge in one area of genetics. Fifth, they can “learn how to learn.”

The fifth experience combines the other four functions. The course is process-oriented, and the goal of the process is to generate content-knowledge, so students learn both process and content. For example, model-revising problem solving helps them learn to take effective action even when they are not sure what to do next in a surprising situation, and it illustrates how to learn a subject deeply, in a way that lets them “think with their knowledge.” And if students become excited about science, or about thinking in general, because they vividly experience “the joy of thinking,” they will be more motivated to want to learn in the future.

I think the Monona Grove genetics course is effective as it is now. But even in high quality instruction there is room for improvement, and the following suggestions might help to improve the course. Of course, for any proposed revision it is wise to ask whether the change would really produce an improvement. I make few claims for the ideas in the remainder of this section, since none of them has been tested in the classroom, but I think, at the very least, that these ideas will stimulate productive thinking about some new possibilities for the course.

A. Supplementing Incomplete or Inauthentic Science Experiences

One goal of my analysis was to determine the similarities and differences between the problem-solving experiences of students and scientists. The analysis indicated one major difference, and several minor differences, between actual research science and the simulated research science that students do in this course.

The major difference is that in science, quick-and-cheap mental experiments help scientists design and evaluate physical experiments that typically require much larger investments of time and money. But with GCK the computer-simulated physical experiments are quicker than a mental experiment, so there is no incentive to do mental experiments. Thus, an indiscriminate doing of simulated experiments is rewarded.[27] This contrasts, for example, with the carefully designed experiments of Mendel. During a growing season, and between seasons, Mendel would analyze the results of the previous sets of experiments, and would design experiments for the next season. If a set of experiments was poorly designed, a great deal of time would be wasted, so Mendel took experimental design seriously, and he was willing to invest a great deal of mental effort in doing it well.

To let students participate in this science experience, perhaps in the MG course there could be a

---

[27]. But during a conference the presenting group feels pressure to design and interpret experiments effectively (and quickly) because the whole class is waiting. The fast pace is described by Lemberger (1995, pp. 273, 275-277): "The pace of the class conferences was always frantic. [The time pressure is evident in the following side-comments by presenters:] ‘K: I'm confused. ... K: Well, we don't have our letters like matched up with this yet, so it's really confusing. L: Yeah. That's true. ... L: If we just had a little bit more time. ... K: All that we know is that it's a 1:1 ratio, but we haven't figured out which letters of those things coordinate with our A, B, C and D. L: I don't know. ... L: I wish she would just give a us a little bit more time here. S: No kidding. (inaudible) K: She's like getting us all mad.’"

supplementary activity whose goal is to determine the characteristics of a field collection (such as the inheritance pattern that is operating, and the genotypes of specific fruitflies) with the fewest experiments. This would be similar to geometry, where a 4-step proof is usually considered better (i.e., more elegant and economic) than an 8-step proof, even if both proofs are equally convincing. Or, for TV watchers, "I can name that tune in two notes" is better than naming it in three notes. Students could evaluate the relative value of the observation-and-interpretation strategies used by different groups, based on criteria such as the number of steps, the time required to reach a conclusion, the accuracy of conclusions, and the sufficiency of experimental evidence — do the proposed theories seem to be a lucky guess, or is there valid evidence to support them? Or is a logical leap that looks like a lucky guess actually a clever and intuitive retroductive inference? Yes, there would be differing opinions about what the criteria should be, and how they should be weighted. But heated intellectual debates about controversial issues would be in keeping with the spirit of the course, and with some of the more interesting episodes in real-world science.

As usual, the problems for this game would have to be carefully designed, in order to be practical and educationally useful. For two reasons — to insure fair competition for the "name that tune" game, and to provide common experiences for comparing experiments and interpretations during discussions — it would be best to give each group the same problem, generated from the same duplicated GCK file. One possible format for the activity is: 1) let each group do a certain number of crosses in a limited time (such as 5 crosses in 15 minutes) in class, 2) students interpret the resultant data as homework, 3) the next day in class they return to experimenting and theorizing, and 4) a conference follows, to discuss and debate the merits of each group's process-and-conclusions. To guard against in-conference adjustments after hearing the observations and interpretations of other groups, each group could submit their conclusions in writing prior to the start of the conference.

Related to this omission is the absence of a need to do sophisticated statistical analysis of data. If a certain experiment produces results that are difficult to interpret, the most efficient strategy is to generate more data (which is quick and easy with GCK) rather than trying to struggle with the question of whether there is satisfactory agreement between existing data and a theory's predictions. GCK does offer a "chi squared analysis" feature to facilitate statistical reasoning, but this feature is rarely used by students. In this case, as above, a supplementary activity could let students experience this important aspect of scientific reasoning.

The rarity of carefully designed experiments is due mainly to the absence of ‘physical experiments’ when students are using GCK. I don't think this is a serious omission, however, since in conventional labs the students get to do plenty of physical experiments, and because there are so many advantages in using the speedy GCK simulations. Besides, in other parts of the course, students do physical experiments. And sometimes Sue brings real fruitflies into the classroom, to let students observe traits and categorize their variations.

With GCK there is also a limited range of experimental design. Students make important decisions about experimental design, but these decisions are limited to asking “Which parents should we cross?” It might be useful, to broaden students' horizons, to supplement their narrow first-hand GCK experience with second-hand stories about the much wider range of experiments (and interpretive strategies) that are possible in a modern genetics laboratory.

During normal research most scientists do not expect their foundational theories to be changed, but during Rounds 1 to 3 the students are told that they will do model-*revising.* A previously existing theory may be selected and then revised, or selected and used as part of a new theory (such as in Round 2 when a new theory is produced by combining their old codominance pattern with a new concept of multiple alleles). Perhaps a new experience — requiring a decision to “select or invent, you figure it out” — could be provided by doing an activity similar to the “name that tune” game described above.

Another omission, the lack of authentic ‘area posing’, was discussed in Section 3.44.

In research science there is no omniscient human who “knows the answer” to a problem, and can supply hints. But one function of the teacher is to adjust problem difficulty so the students' experience is *more* like that of research scientists. And often a project leader serves a similar supervisory role by coordinating the efforts of scientists in a research group.

B. Using ISM in Discussions of Problem-Solving Strategies

As discussed in Section 3.51, the teacher suggests that the problem-solving skills of students could be improved by holding 'strategy discussions' between rounds. This is consistent with the suggestion of Finkel (1993) to teach model-testing strategies, and with instructional techniques that are designed to help students learn more from their experience. A potential benefit of explicitly calling attention to relevant aspects of experience is described by one of the course developers:

> Whether or not the solver recognized that the genetics-specific heuristics were all examples of more general heuristics, and as a result would have been able to apply the more general ones in new contexts, is uncertain. It is unlikely that they would be recognized unless an instructor made them explicit. ... If students are made aware of these general procedures, so that they develop an abstraction of the procedure separated from the specific genetics content, then the possibility for transfer is increased. (Stewart, 1988, pp. 251, 240)

This suggestion is consistent with the concept of 'forward-reaching transfer' proposed by Salomon & Perkins (1989).

One aspect of an effective teaching environment, designed with the goal of improving the "experience → knowledge" connection, is to increase students' metacognitive awareness of what they are doing and of "what they can learn" that will be useful in other contexts. To help students increase their overall awareness, and to call attention to specific opportunities for transfer, discussions of general heuristics and their domain-specific applications (as described in Stewart, 1988; and in Stewart, et al, 1992, and Hafner & Stewart, 1995) could be a valuable component of effective instruction.

I enthusiastically agree with Johnson, Finkel, Stewart, Salomon and Perkins, in their recommendations for explicitly helping students to be more aware of what they are doing and of what is happening around them. I think it would be useful to discuss 'Darden strategies' (as suggested by Johnson, 1996) with students, and also genetics-specific and general heuristics (Stewart, 1988). But it might be even more useful to discuss a combination of "strategies + heuristics + ISM," and I think there could be a smooth blending of these three.

In this blend, ISM would provide a comprehensive overview of the 17 science experiences and how they work together in functional synergy during problem-solving research. Specific strategies and heuristics would be used in detailed 'blowups' that focus on particular regions of ISM, describing them in more detail than is contained in the basic ISM framework. Some principles for using ISM in the classroom, such as "intentional learning" and strategies for "coping with complexity," are discussed in Sections 4.21-4.22.

C. Using Prediction Overviews

During my analysis of the MG classroom, I invented a new tool for thinking about genetics. This tool, a 'Prediction Overview',[28] is a visual representation that shows the predicted cause-effect relationships in genotype-to-phenotype mappings, how "all possible results viewed at the level of genotypes" cause "all possible results viewed at the level of phenotypes." A Prediction Overview (PO) is especially helpful for working with a new model — for understanding its genotype-phenotype relationships, making predictions, and designing experiments. This understanding, and these skills, are essential for the GCK model revising that is the central activity in this course.

POs could help students build a solid foundation of 'working knowledge' that fluently combines content knowledge (of genetics) with procedural knowledge (of problem solving). A good working knowledge of genetics is valuable in all aspects of problem solving — for making effective decisions regarding experimental design, data interpretation, theory invention and development, and empirical and conceptual evaluations.

POs could be a valuable part of the enriched discussion of genetics recommended by the teacher. But to maintain an appropriate level of difficulty, I don't think POs should be introduced until after Round 1, because a PO would make Round 1 too easy to solve. But in Round 2 the use of POs — if the process of making a PO is clearly explained by showing the principles in a step-by-step process, so students can apply the PO-principles for new models — would make it easier for students to cope with the complexity inherent in Round 2, without giving away the basic concepts involved in a solution.[29] POs would also be useful for Rounds 3 and 4.

It is important to recognize that a PO and Mendel's Bible (MB) are complementary, not competitive. Both representations provide a useful overview of a model, so using both representations will allow a model to be viewed from two different yet related perspectives. Each

---

[28]. Prediction Overviews are described in detail in Appendices C20-C24.

[29]. It is even possible, as explained in Appendix B15, to invent a PO to make predictions for the 'three alleles per individual' models that are often proposed during Round 2.

type of representation is useful for different purposes; usually a PO is better for making predictions based on a new model, and for interpreting data, while MB is better for summarizing the components of existing models, and for 'invention by revision'. Therefore, my claim is not that a PO is superior to MB, but that a combination of both is better than Mendel's Bible by itself.

So far, however, all of this is speculation. Because POs have not yet been used in the classroom, there is no empirical data on how the use of POs would affect the process of problem solving and the learning of conceptual and procedural knowledge. But I think that Prediction Overviews have significant potential for being pedagogically useful.

3.6: Evaluating the ISM-Based Analysis

This section discusses the extent to which the ISM-based analysis of the MG course has achieved the goals set for it. Chapter 1 describes two sets of goals, which are examined in the two sections that follow.

3.61: Understanding the Structure of Instruction

In Chapter 1, the first set of goals focuses on the instruction:

> The immediate goal of analysis is to construct a multi-faceted representation of the instructional methods, student activities, and student actions in one classroom, with the purpose of gaining a deeper understanding of the functional relationships among various activities and actions, and of the ways in which student actions are related to the actions of scientists.

To what extent has my analysis succeeded in accomplishing these goals? In the following paragraphs, this question will be discussed one phrase at a time. Two definitions from Section 3.22 will be used; a 'student activity' is what students are asked to do, and 'student actions' are what students do in response to the activity-request.

Multi-faceted representation? Sections 3.23-3.48 contain an analysis-based description of student activities, student actions, and instructional methods. Yes, I think this description is multi-faceted, examining many characteristics of the course from different perspectives.

Student activities and student actions? My analysis doesn't add much to the description of student activities by Johnson (1996) in Section 3.23; in the A-and-E grids in Tables 3-7 these student activities are simply listed at the top of each column. But there is a discussion of another major activity (listening to stories) and two sub-activities (interactions with other students and with the teacher) in Sections 3.24 and 3.45-3.47. And the analysis does contribute to a detailed examination of student actions, which are summarized in the 17 ‘science experience’ cells in each column, and are discussed in Sections 3.31-3.35 and 3.48_A.

Instructional methods and functional relationships? The instructional methods used in the MG course, including planned activities and improvised actions by the teacher, are the focus of Sections 3.23-3.25 and 3.41-3.48. The functional relationships within individual activities are discussed in Section 3.48_A, and the functional relationships among activities are examined in Sections 3.41-3.48. I think the description of functional relationships is accurate and reasonably thorough.

A deeper understanding? Personally, I am pleased with the understanding of the MG classroom that I have gained during the analysis. I hope my discussions have helped the reader to also gain an improved understanding of the course.

Student actions and scientist actions? As explained in Sections 3.22 and 3.26, there is a close connection between scientist actions, ISM, and the A-and-E grids used in my analysis. Therefore, whenever a student action is discussed in an A-and-E cell, students are doing approximately the same type of action as in actual science. But is the students' simulated research similar to “the real thing”? Most aspects of the student experiences seem fairly authentic, especially when first-hand and second-hand experiences are combined. A few exceptions, where student actions differ noticeably from the typical actions of scientists, were discussed in Section 3.52_A.

In using ISM to evaluate the relationships between student actions and scientist actions, there is a ‘transitive’ character to the logic that is used. In mathematics, transitive logic occurs in a situation where A and B are both equal to C, and the transitive conclusion is that A and B are equal to each other. Similarly, student actions and scientist actions can both be described in terms of the ISM framework, which serves as a common basis that makes it easier to compare the classroom activities of students with the research activities of scientists, to determine the ways in which they are similar and different. This transitive comparison can contribute to an understanding of the ways in which one of the original goals of the course designers — to let students experience the activities of authentic scientific research — has been achieved. One "transitive logic" viewpoint, which can be useful if applied with insight and caution, is that because the essential actions of scientists are summarized in ISM, when students are doing the actions in ISM they are doing the actions in science.

3.62: Testing and Improving the Analytical Utility of ISM

The first set of goals focused on the MG classroom. A second set of goals involves ISM:

> Viewed from a long-term perspective, the goal is to use this analysis as an opportunity for testing the analytical utility of ISM, gaining a deeper understanding of ISM, and improving its effectiveness as a descriptive framework and as an analytical thinking tool that can be used, by curriculum designers and teachers, for developing instructional methods that will expand the range of learning experiences for students.

Have these goals been achieved? Sometimes, in some ways.

A. Testing ISM as a Tool for Instructional Analysis?

The first goal in this set is to test the analytical utility of ISM. One way to do a test is to run a thought-experiment as a 'control' by imagining what the analysis might look like without ISM, and comparing this result with the analysis as it is. When doing this, my conclusions produce a "mixed review" for the analytical utility of ISM.

For the first phase of analysis, the result with ISM is definitely better than it would have been without ISM. According to my thought experiments, the vertical relationships in the A-and-E grids, showing the 'multiple experiences' within each activity, have been especially improved by using ISM. But what is the reason? Without ISM I would not have invested nearly as much time in the analysis, because each of the 204 cells in the five A-and-E grids forced me to think about whether (and how) each classroom activity involved students in each type of science experience. So is the analysis improved due to ISM (via the A-and-E grid) or because of the time I invested? Could I have achieved the same result, or better, without ISM? Probably not. Why? First, the many questions posed by an initially empty A-and-E grid helped to stimulate productive thinking about student experiences. Second, the structure of the grid, and the structure of the ISM framework on which it is based, helped to organize the data, thus making it easier to recognize 'function in science' patterns in the students' actions. Third, the analysis made it easier to see specific ways in which the students' experience differed from the typical experience of research scientists. Fourth, the analysis required an understanding of science, and much of what I know about science is a result of what I have learned by developing ISM.

For the second phase of analysis, the answer is easier. It seems that most of my 'repeated experience' conclusions could have been reached without using ISM. Certainly someone who had invested a significant amount of time studying the course (or for some conclusions, just a little time) could have made similar conclusions. So is ISM without value for studying the 'repeated experiences' that appear horizontally on an A-and-E grid? No; I think ISM is valuable for analysis. One reason for this utility — the visually meaningful organization of information in a grid, which can promote an improved understanding of pedagogically functional relationships — was discussed in Section 3.21. During my analysis, the structure of the grids did stimulate thinking about repeated experiences, did require making distinctions between the 17 science experiences and searching for functional relationships involving each experience, and did make it easier to literally "see" what is happening in the course. This was especially helpful in the second phase of the analysis.

If ISM was helpful in doing the analysis, then why would the results of analysis have been similar, with or without ISM? This is mainly due to the structure of the MG course. The central activity, GCK model revising, is fairly complete and self-contained, providing an opportunity to experience almost the whole range of experiences that occur in research science. And it is relatively easy to see what the students do — i.e., almost everything — even without using ISM. The Black Box model revising is obviously analogous to GCK, so the parallels (and the forward-looking preparation that links the Black Box with GCK) are easy to recognize. The function of the content-constructing activities is also easy to see. And if anyone missed the 'persuasion' function of the conferences and manuscript preparation, a quick reading of papers about the course, written by the course designers, would make it clear how the last of the 3Ps is manifested in the classroom. But for another course that was designed in a way such that the functional connections between activities were less clear, I think ISM would be more useful because it would be more necessary. There seems to be a 'ceiling effect' with the MG course, because the analysis of this clearly structured classroom was not sufficiently challenging to properly test the usefulness of ISM and the A-and-E grids.

In another setting, in which the structure of the original design is less clear, ISM could be very useful for curriculum analysis and revision. One such use of ISM is discussed in Section 4.11, using an example from my experience in teaching chemistry labs. Another potential analytical application involves the use of ISM for generalizing educational principles from one instructional context to another. This type of application is discussed in Section 4.12.

B. An Improved Understanding of ISM-Based Analysis?

The second goal (in the second set of goals) asks whether the analysis was useful for "gaining a deeper understanding of ISM." Here, again the answer is easy, but now it is “yes!” Specifically, the process of analysis led to a recognition of some ways in which ISM was incomplete because an element was either missing or was not clearly characterized, or because an integrated relationship was not included in the ISM framework or in its elaboration. These recognitions of flaws have led to many changes, major and minor, in the ISM framework and its elaboration. Two major revisions inspired by the MG analysis were a change in terminology from the initial ‘degree of surprise’ to the more intuitive ‘degree of predictive contrast’, and making a distinction between domain-theories and system-theories. In addition, the challenge of trying to apply the 3Ps to the MG course motivated me to clarify the ISM definitions for these terms and for other concepts such as action evaluation, levels of problem solving, problem, project, and preparation. Even for the parts of ISM that have not been changed, my understanding has been deepened by the instructional analysis, especially by my in-depth analysis of GCK model revising.

C. An Improvement in ISM as a Tool for Analysis?

The third goal asks whether the MG analysis has improved ISM's "effectiveness as a thinking tool" for use in curriculum analysis and design. Due to my experience with the analysis, I am certainly more adept at using ISM for analysis. Some of this knowledge is personal, and would not transfer to another analyst using ISM or the A-and-E grid. But some of the results of this experience could be formulated into principles that could be communicated to others who want to use ISM for curriculum design. In addition, the content and format of the A-and-E grid has been improved since the original prototype in 1994, partly because ISM has been improved due to my experience in analyzing the MG course.

One improvement is simply having a grid, which would not exist if the MG analysis had not necessitated it. The ISM diagram and grid are similar in some ways, but different in others. For example, the diagram (with its free-form spatial arrangements, plus a symbolism that includes connecting lines and arrows,...) is better than the grid (with its simple columns and rows) for expressing the integrated relationships in science. But the grid, with its temporal dimension showing the sequencing of activities (Activity #1 is followed by #2 and #3 and...), can stimulate insights that are not as easy to perceive when using the diagram. More important, the grid makes analysis possible. It is difficult to imagine how Table 8, with its coordination of 11 instructional activities, could be done using the ISM diagram. In the grid there is a grouping of science experiences into seven categories, which can be useful as an organizer of thinking. And even though there is some loss of visual integration, compared with the diagram, this has not been a serious problem for me. While working with the grid, I have noticed that the 'science experience' labels (in the first column of the grid) are sufficient to remind me of the integrated relationships that are now a part of my mental models for science. The ISM diagram played an important role in constructing these vivid mental models, so the diagram exerts an influence even when I am just using the grid. Because elements of scientific thinking are logically organized according to integrated relationships in the ISM framework (especially in the diagram), it can be easier (even when using the grid) to find the 'function in science' relationships between diverse elements of classroom experience.

D. Using ISM as part of an Eclectic Analytical Framework?

Another potential application for ISM involves its use in an analytical framework constructed by eclectically combining ISM with ideas from one or more complementary perspectives.

One category of potential contributors includes frameworks that have been used in previous studies of GCK problem solving and/or the MG course. As described in Appendix B20, GCK problem solving has been studied in seven doctoral dissertations done at the University of Wisconsin-Madison. Each of these studies examined the classroom using a different interpretive framework: Collins (1986) used the problem-solving ideas of Reif (1983), Thomson (1993) used his own modification of ideas borrowed from Darden (1991), Hafner (1991) used his own "Model-Revising Problem Solving in Genetics" framework (adapted from Klahr and Dunbar, 1988), Finkel (1993) used several frameworks, including three types of knowledge and a list of ten descriptors that she developed during her analysis, Wynne (1995) used a framework that included ideas about theory evaluation from Clement (1989) and Lakatos (1978), Lemberger (1995) used the Conceptual Change Model (Posner, et al, 1982), and Johnson (1996) used the problem-solving strategies of Darden (1991). Each of these analyses has value, and much can be learned by comparing the insights gained by those who have viewed the process and results of teaching and learning from different perspectives.

It might be possible to construct a useful framework by combining ISM with concepts from one or more alternative frameworks — either a framework that has been used to study GCK/MG, and is listed above, or one that is not listed. For example, ISM contains very little detail about the cognitive processes involved in problem solving, so a useful eclectic framework might be constructed by combining the basic concepts in ISM with a more detailed psychological model from cognitive science. Another component of a useful eclectic mix might be ‘thinking strategies’ such as the general heuristics and content-specific heuristics and algorithms described by Stewart (1988). Another potential partner is the framework of Perkins & Simmons (1988) with its 4 frames of knowledge (content, problem solving, epistemic, and inquiry) and 3 forms of knowledge (strategic, autoregulative, and beliefs); the frames run parallel to the 17 science experiences in ISM, describing the same types of actions, and the forms are orthogonal to the frames and to ISM. These possibilities, and others, may be worth investigating in the future.

A discussion of potential analytical and educational applications for ISM continues in Chapter 4.

****

CHAPTER 4

Potential Educational Applications for a Model of 'Integrated Scientific Method'

This chapter explores some ways that a model of 'integrated scientific method' (ISM) might be used in education. I will discuss *potential* educational applications of ISM, but will not develop specific applications. Thus, there will be a discussion of possibilities for future development, rather than a description of applications that already have been developed. Potential educational applications for ISM could occur in three ways: by using ISM to facilitate the design of instruction, to teach students in the classroom, or to educate teachers. These potential applications will be discussed in Sections 4.1, 4.2, and 4.3, respectively. Section 4.4 explores some possibilities for applying 'thinking methods' to a wider domain of knowledge, including science and much more. The dissertation concludes with a brief overview in Section 4.5.

The following discussion builds on the foundation outlined in Section 2.00. One essential assumption is that students will benefit, cognitively and/or culturally, if they learn about the process of science, to supplement their learning of science content. Another assertion, which is the focus of discussion in this chapter, is that learning about "the process of science" might be improved by using ISM to describe this process.

4.1: Using ISM for Instructional Design

Section 4.11 examines the use of ISM to stimulate the development of 'Aesop's Activities' that are intended to fulfill specific instructional functions. Section 4.12 discusses the relationships between the analysis and design of instruction.

4.11: Aesop's Activities

An important part of the overall design process is the selection or invention of student activities. To focus attention on the principle that instruction should be goal-directed, that instructional activities should be done for a purpose, a useful metaphor (from Rusbult, 1989) is based on analogy to Aesop's Fables. Just as Aesop's Fables are designed to teach specific lessons about life, Aesop's Problems (for the purpose of promoting first-hand experience in doing science) and Aesop's Stories (for second-hand experience about science and scientists) can be designed to help students learn lessons about the process of problem solving and the nature of science. An activity-and-experience grid, with ‘science experiences’ based on ISM, could facilitate the design of Aesop's Activities (both problems and stories) by stimulating and structuring a search for activities to help students learn more about each element in ISM, either by itself or combined with other elements into coherent activity-modules that teach the integrated relationships between different science skills. The visual organization of information in a grid, to show functional relationships within and between activities, can make it easier to see opportunities for the sequencing and overlapping of activities within a course.

The usual objective of instructional design, of course, is to select or invent activities that are pedagogically effective. **A suggested use of ISM to orient the design of Aesop's Activities is based on an assumption that** doing what scientists do (for first-hand experience) or learning about what scientists do (by second-hand experience) **will help students understand the process of science. If this assumption is valid, then it is reasonable to expect that activities based on ISM** (which are the activities of scientists) **could be educationally useful**. But this expectation will not be considered a proof. Compatibility with ISM is certainly not the most important indicator of probable pedagogical effectiveness. Instead, a wide range of relevant factors should be considered, including student motivations and learning styles, the cognitive and emotional maturity of students, and limitations on teaching time. To supplement a conceptual analysis of these influencing factors and their interactions, useful information might also be gathered by empirically testing the effectiveness of instruction; informative experiments to examine learning outcomes could be done at the level of an activity, a course, or a curriculum. But this type of testing will not be a part of my dissertation.

However, I have done a small pilot experiment to test the effectiveness of ISM in stimulating and structuring ideas for Aesop's Activities. For the past 11 semesters I have been a Teaching Assistant for introductory chemistry courses at the University of Wisconsin. The labs associated with these courses are coordinated fairly well with the course lectures, but not with each other. In an attempt to add more coherence to the set of labs, one semester (Fall 1995) I used ISM to ask, for each lab, "Which science actions are students doing?" At the beginning of the course, I analyzed labs for the whole semester and decided which aspects of the students' experience would be worth calling attention to (by discussions or small assignments) during each lab. I was not willing to invest a large amount of effort in this analysis, but based on my estimates of the 'return per hour invested' I think that ISM was very useful in stimulating ideas and planning a structure for the instruction. This contrasts with the MG course where (as discussed in Section 3.62_A) the instruction is already so well organized that the structure is easy to see, with or without ISM. My experience in improving the coherence of the chemistry labs, which had less structure initially, gives me confidence that ISM will be useful for the analysis and development (by generation or revision) of instructional activities.

4.12: Analysis and Design

The use of ISM for instructional design — which may involve either curriculum policy decisions regarding existing instructional programs, or the design of new programs — can be viewed as an extension of using ISM for instructional analysis, as in Chapter 3. When making any decision about instruction, whether this involves a program that is old or new, the starting point should be an accurate understanding of the program's characteristics. This understanding, which can be improved by analysis, will provide a basis for evaluating the ways in which the program does or does not meet a set of objectives, for envisioning the types of revision that might help it to better meet these objectives, and for making decisions about revision. Because analysis can facilitate the understanding that is essential for design, a tool that is useful for instructional analysis may also be useful for instructional design.

When activity-and-experience grids (based on ISM) are used to interpret what occurs in the classroom, this can facilitate an improved understanding of student experiences within each activity, and of the functional relationships between student activities. One important type of relationship involves the ways in which activities are selected and sequenced so that students will have repeated experiences with each skill, many times during a course, in different contexts and at increasing levels of sophistication, in a ‘spiral’ that promotes improved understanding and proficiency. Based on an understanding, gained from ISM-based analysis, of how ‘repeated experiences’ are used in an old course, a new course could be designed with the objective of using these principles to improve the synergism between mutually supportive activities. In this way, knowledge gained from analysis can be used to generalize instructional methods and principles from one context to another.

For example, an ISM-based analysis could serve a useful function in applying principles from the MG course to modify another course, in a five-step process. First analyze the MG course using ISM, as in Chapter 3. Second, similarly analyze another course (the one that will be modified) to characterize its content-and-structure. Third, compare the results from Steps 1 and 2, to search for insights about similarities and differences between the courses; because the same methods have been used for each analysis, this common basis should facilitate the comparison. Fourth, creatively think about possibilities for modifying the other course, with the goal of incorporating some principles from the MG course in order to achieve the objectives set for the other course. Fifth, evaluate the options, to decide which ones might be productive. In this type of application, with the objective of generalizing instructional principles from one setting to another, an ISM-based analysis could promote a deeper understanding of the structure in each course, and of the possibilities for transferring methods (suitably adapted) from one course to the other.

Another example, described earlier, is the use of analysis-based design to guide a search for goal-oriented Aesop's Activities. This example and the one above (for generalizing MG principles) can be viewed as variations on a strategy for analysis-and-design that adopts a problem-solving approach to ‘invention by revision’: First, use analysis to understand the content and structure of the instruction that already exists. Second, define a set of objectives for what the instruction is intended to accomplish. Third, try to estimate the ways in which the instruction does and does not achieve these objectives. Fourth, compare the perceived now-state (in Step 3) with the imagined goal-state (in Step 2); the difference between these defines a problem. Fifth, think about ways to modify the current instruction, in an effort to bring it closer to a goal-state that can achieve the desired objectives. Additional steps can occur by supplementing this conceptual evaluation with observations of student response to the instruction (in its original form, and later in modified form) to see what students are doing and what they are learning. This empirical information can help to form an accurate characterization, in Steps 1 and 3, of the current state of instruction.

The second step above is important, and is worthy of special attention. Even though it can sometimes be valuable, as in the MG genetics course, to let students experience a wide range of scientific methods, "doing everything in ISM" is not always a wise goal. When teaching time is limited (and this usually seems to be the case), instead of trying to do everything (and doing it at a fairly shallow level), it may be more effective to focus on doing a few selected aspects well. Or, a curriculum designer or teacher may decide that, for some students at a certain stage of their education, a simpler model of science will be more useful. For example, I think that all students would benefit from understanding the basic principles of hypothetico-deductive inference, and that the dual-parallel 'box' in ISM — with the same process being done mentally (to produce predictions) and physically (to produce observations) — is an excellent way to think about this process of inference. And there are many possibilities for transfer of knowledge and skills, because a similar type of parallel logic is also used in retroduction and in the generalized 'integrated design method' that is shown in Figure 9 and is discussed in Section 4.41. But even if this aspect of ISM is considered essential, in some educational contexts it may be wise to avoid using other aspects of ISM, either for practical reasons such as a limited amount of instructional time, or in order to simplify the models that students are asked to learn.

Another important consideration is the optimal level of "realism" during simulated science in the classroom. For helping students learn, teaching experience and/or formal research may indicate that a modified version of science process seems to be more pedagogically effective than an "authentic" mimicry of what scientists actually do. But even if modified versions of scientific methods are used in the classroom, ISM can still be useful in providing a baseline foundation for understanding the methods of science, and for stimulating ideas about how these methods might be modified to produce effective instruction.

4.2: Using ISM in the Classroom

Section 4.1 discussed the potential value of using ISM for the goal-oriented design of activities that promote a wider variety of student experiences, and a more effective blending of experiences. Section 4.21 will focus on the direct use of ISM in the classroom to help students *learn* from their experience — and *remember* what they have learned, and *transfer* this knowledge to new situations — by using ISM to explicitly direct attention to important aspects of "what can be learned" from each type of experience, and by providing a coherent, meaningful framework for what is being learned. Section 4.22, "Coping with Complexity," will discuss the unavoidable tension between completeness and simplicity, and how the pedagogical effectiveness of ISM might be improved by the flexible adjustment of key instructional factors such as information content, and by carefully planned pacing and sequencing. Section 4.23, "Should Scientific Method be X-Rated?", discusses 'controversial issues', focusing on issues that are especially relevant for education.

4.21: Learning from Experience

My dissertation did not include any empirical testing of classroom applications for ISM. Instead of gathering data on which to base an empirical evaluation, I will attempt to show that — based on principles of learning-and-teaching proposed by other educators (such as Perkins & Simmons, 1988; Perkins & Salomon, 1988; Bereiter & Scardamalia, 1988; Mayer, 1993) — it is reasonable to expect that ISM could have significant pedagogical value. A brief introduction to these principles follows, beginning with a model that is closely related to the methods of science, and thus to ISM.

Most modern educators agree that high-quality education should help students master both content and thinking skills. For example, *Science for All Americans,* published by Project 2061, states that "knowledge should be understood in ways that will enable it to be used in solving problems. In this sense, all of the foregoing recommendations [about content-knowledge] are about thinking skills. (Rutherford & Ahlgren, 1990, p. 175)" In an effort to describe the dynamic relationships between thinking skills and the learning of subject-area content, and to encourage the development of instructional techniques that will help students develop a deep understanding of content and thinking skills, Perkins & Simmons (1988) propose an integrative model with four mutually interactive frames of knowledge: content, problem solving, epistemic, and inquiry. After describing each frame in detail, along with examples that illustrate the detrimental effects of ignoring some frames during instruction, or of treating the frames in isolation from each other, the authors — based on their theory that "people learn much of what they have a direct opportunity and some motivation to learn, and little else" — recommend that "instruction should include all four frames ... and should involve explicit articulation by teachers and/or students of the substance of the frames and their interrelationships. (Perkins & Simmons, 1988, pp. 319, 321)"

To explain why a four-frame articulation should be explicit, Perkins and Simmons simply state that "direct opportunity to learn about the frames and their interrelationships is fostered by explicit treatment. (p. 321)" This explanation is supplemented in other papers (Perkins & Salomon, 1988; Salomon & Perkins, 1989) where Perkins claims that transfer-of-learning can occur by way of a 'high road, forward-reaching' cognitive mechanism in which basic elements of what is being learned are consciously abstracted and decontextualized from one specific context in anticipation of their later application in other contexts. This mechanism can be pictured in terms of a 'bridge' analogy; if an idea is linked, in the mind of a learner, to a number of different contexts — including future contexts that are being imagined — this idea will form a bridge that promotes transfer from one context to the next.

An explicit, conscious focus on "what can be learned" is also consistent with a theory of *intentional learning* (Bereiter & Scardamalia, 1988) which claims that students will learn more if they invest extra mental effort, over and above what is required merely to accomplish schoolwork tasks, with the intention of pursuing their own cognitive goals. The goal of intentional learning is to transform a current state of knowledge (or level of skill) into an improved future state (or level). Thus, according to ISM's definition of *problem solving* as the process of transforming a current now-state into a future goal-state, intentional learning is *a problem-solving approach* to learning. An effective strategy for achieving intentional learning combines an introspective access to the current state of one's own knowledge, the foresight to envision a potentially useful state of knowledge that one currently does not possess, a plan for how to transform the now-state into the goal-state, and a motivated willingness to invest the time and effort that will be required to reach this goal.

Forward-reaching transfer and intentional learning are closely related, and both strategies are activated when a student wisely asks, “What can I learn now that will help me in the future?” The future value of “what is learned now” may involve content or process, or both. For example, the principle of energy transformation is an essential component of content across a wide range of subject areas, so learning about energy in one area should facilitate learning about energy in another area. Similarly, thinking skills that are useful for the process of problem solving in one field of science will often be useful in other areas, both in and out of science. Some relationships between content and process are explicitly characterized in the Perkins-Simmons model with its four interactive frames of knowledge. In science these frames can be described in terms of thinking skills: the *content* frame is **learning** scientific theories, *problem solving* involves **using** these theories, the *epistemic* frame is **evaluating** theories, and the focus of *inquiry* is **inventing** theories. With this formulation, one way to explicitly articulate "the substance of the frames and their interrelationships" — and to pursue the educational benefits that may ensue — is to show how the frames operate in the context of science. This is what my model of ‘integrated scientific method’ does.

As a way to articulate the four frames, ISM offers two distinct benefits. First, the visual organization of ISM makes it easier for students to understand essential relational patterns, to literally see how details fit into the ‘big picture’ of science, because an explicit, logically organized *visual model* can help students construct their own *mental models* of science and thinking strategies (Mayer, 1993). Second, although it is closely related to the four-frame model, scientific method is more familiar to scientists, study-of-science scholars, educators, teachers, and students, so it may be easier to communicate ideas effectively if they are expressed in terms of scientific method, used by itself or in conjunction with the four frames of knowledge. This familiarity will also make it easier to establish connections with the large amount of thinking that has been done about the methods of science and their application to education.

4.22: Coping with Complexity

To be effective, a teaching technique must aim at an appropriate level. In every model, including ISM, there is a tension between the conflicting criteria of completeness and simplicity. Generally, with a model of science, increasing the complexity of a model (if this is done skillfully) will improve its potential for accurately describing science, but if a model is too complex there will be difficulties in making it pedagogically useful. For example, if too many concepts in ISM are presented too quickly, students may become overwhelmed and confused. For any type of instruction, effective teaching requires wise decisions about the selection, pacing, and sequencing of what will be taught. In constructing ISM, I have tried to facilitate these decisions in two ways — by aiming for an intermediate level of complexity that is reasonably close to levels that will be useful for typical instructional applications, and by using a flexible ISM framework that lets the complexity be adjusted by using the techniques of *simplification* or *supplementation* described in this section.

To achieve a useful balance between completeness and simplicity, so that ISM can be used for students of different ages, abilities and interests, adjustment mechanisms are necessary. The simplest mechanism is to adjust the *information content* of ISM by describing it in *less detail* than in the framework or in my elaboration, or to describe some parts of ISM in *more detail* by supplementing these parts with ideas from other sources. In either case, an extensive use of concrete illustrative examples will help students translate the abstract concepts of ISM into meaningful personal knowledge. Of course, as with any other subject, instruction whose goal is to improve students' ability to solve problems or to understand “the nature of science” will be more effective if there is an appropriate level of difficulty at all times, with well designed activities that are wisely paced and effectively sequenced.

One mechanism for visual simplification is to ignore parts of the ISM diagram, either permanently (by removing some elements from the model) or temporarily (by showing, in an *isolation,* only a part of ISM, a part that students have experienced recently and that currently is being discussed). Isolations can be combined with *blowups* that show, in more detail, the part being studied. Isolations and blowups will be more effective if they are used in a coordinated plan of *pacing and sequencing,* such as a whole-part-whole teaching method that shifts back and forth between the whole ISM diagram and isolations (with or without blowups). Used skillfully, this teaching strategy will help students learn more about each part of ISM and its relationship to other parts and to the whole.

This method, when used in contexts of gradually increasing complexity, can also help students improve their ability to "cope with complexity" by developing the mental skills that are required to understand complex systems of concepts. Another aspect of thinking that will improve with practice is the ability to interpret diagrams. The organized complexity of the ISM diagram provides opportunities for practicing and improving valuable visual skills that, in conventional education, are typically under-used and thus underdeveloped. But the main goal of the organized complexity in ISM is to improve conceptual understanding by stimulating a viewer to see relationships in new ways and thus to gain a deeper understanding of science.

A complementary combination of verbal and visual representations, as in ISM, can be more effective than either format by itself. An explicit, logically organized *visual model* such as the ISM diagram, which uses spatial relationships to show conceptual relationships, can help students recognize essential relational patterns and construct their own *mental models* of scientific methods and problem-solving strategies. Winn (1987) explains,

> Diagrams are effective in instruction because they allow students to use alternative systems of logic. ... Presenting information graphically allows students to scan it rapidly to discover patterns of elements within the diagram that are meaningful and that lead to the completion of a variety of cognitive tasks. ... [In this way] certain physiological strengths of learners, such as pattern recognition,...can be exploited. (pp. 157, 160)

4.23: Should Scientific Method be X-Rated?

This section is a continuation of Sections 2.44-2.45, which examine controversial issues that have sparked heated debates between advocates of contrasting and apparently irreconcilable positions. This section does not try to provide answers. Instead, it will discuss issues, express opinions, and make modest recommendations based on a simple principle: If a good idea is taken to extremes, without sufficient balance from critical thinking and common sense, there may be undesirable consequences.

Although ISM is intended to be descriptive, not prescriptive, when any model of science is used in the classroom there will be a normative influence on students. In fact, promoting a change in student ideas and behavior is the purpose of education, and is the main motivation for including any model of science in a curriculum. Therefore, when selecting a model of science to be used for education (in the classroom or for curriculum design), two important questions should be asked: What is the most accurate description of science, and what educational approach is most beneficial for students?

What is the opinion of the ISM framework regarding major controversial issues? Usually, “no comment.” ISM (i.e., the framework) seems to be relatively neutral, although of course it cannot be totally neutral. By contrast, my ISM elaboration contains many weak opinions and a few strong opinions. But my elaboration of ISM is only one of many possible alternative elaborations; the same ISM framework can be used to express a wide variety of divergent views about science, as illustrated in Section 2.08.

ISM is designed to be flexible, so that by varying the definitions, emphases, and interrelationships of various elements, it can be used to describe a wide range of views about science and scientists. In my opinion, this is a strength, but it is also a cause for concern. If wild ideas are expressed using ISM, should I feel responsible, despite my disclaimer that “the opinions expressed using ISM are those of the expresser, and not necessarily those of ISM”? One area of ISM is especially prone to misunderstandings and mis-use. The potential trouble spot is "cultural-personal factors." These factors are a part of science, so I defend their inclusion in ISM, and a strong case can be made for including them in education. As usual, however, if this basically good idea is taken to extremes, with exaggerated interpretations, the result will be a distorted picture of science. And the ISM framework can be used to express extreme views.

This important issue is addressed by Stephen Brush (1974), who asks a serious question in a humorous title, "Should the History of Science Be Rated X?" Why is he asking this question? Because, as explained in a subtitle, "The way scientists behave (according to historians) might not be a good model for students." In terms of the “accurate and beneficial” questions framed above, we can ask whether "the way scientists behave (according to historians)" is the way scientists really behave; and if they do, are students better off not knowing? These questions are relevant for ISM because it claims that ‘cultural-personal factors’ and ‘thought styles’ affect the process and content of science. Pedagogical considerations of “what is beneficial” should be heavily influenced by, but not totally determined by, what is regarded as most accurate; the possible effects on students and society should also be considered.

Unless science education is totally ineffective, it will affect students' attitudes toward science and toward life. The effects of education can be either harmful or beneficial. But there are debates about effects. For example, will the enthusiasm of future scientists be dimmed if their role models are tarnished by cynical portrayals of scientists as politically motivated, status seeking, self promoting mercenaries? Or will some students want to be scientists because they see the socially interactive aspects of science, and they realize that scientists are real humans, like themselves? Similar questions can be asked about extreme logical skepticism. Should students stop doing experiments because observations are inevitably biased and unreliable? And will students stop trying to learn the theories in textbooks if they are told that these theories are unreliable — or, with anti-realist interpretations, that science does not claim to (or even attempt to) represent the truth about nature? Or will skepticism encourage healthy critical thinking? And is there a danger when science education becomes politicized so that it argues for certain metaphysical or ideological views? Or might certain types of politicization be beneficial for students?

Open minded tolerance for a variety of views can be a virtue, but an educator should be willing to "draw the line" at a point where the views (and arguments used to support them) begin to cross over into areas that seem foolish and even dangerous. In my opinion, some scholars in the 'study of science' community have crossed over this line. Way over. For example, do scientists really create reality?

> It is not that we might come to learn of the molecular weight by means of the apparatus, but it is somehow actually brought into existence by the equipment. Without the material environment of the laboratory none of the objects could be said to exist. (Latour & Woolgar, 1979, p. 69)

Radical sociology of science is an extreme view that, in my opinion, deserves harsh criticism. And this is what it gets in Appendix $A24_B$. A few additional samples from Slezak's strong critique of the 'strong program' follow, beginning with one of the more colorful and questionable aspects, the embracing of postmodern deconstruction by some authors:

> [Latour and Woolgar (1986, p. 273) claim that their] own work, then, just like all of science, has no determinate meaning since, as they explain, "It is the reader who writes the text." ... This move, a characteristic feature of Deconstructionist writings, has the effect of showing any

> critic to have *ipso facto* failed to understand the subtlety and sophistication of the work. In particular, the use of logic and the traditional categories of thought is sufficient evidence of a critic's misunderstanding. (Slezak, 1994, p. 332)

In important ways, radical sociology can undermine the conventional view that "a central aim of education...is the fostering of rationality, or its educational cognate, critical thinking. (Siegel, 1989, p. 21)" By contrast,

> Traditional conceptions of both the ends and the means of a science education are gravely challenged if social constructivist doctrines are taken seriously. If, as Bloor (1976) suggests, beliefs are not a matter of reasons, evidence and other rational considerations, then teaching cannot involve conveying ideas through *understanding*. (Slezak, 1994, pp. 267, 354)

According to Slezak, the pernicious societal effects of radical relativism extend far beyond the realms of scholarly discourse where they originate:

> My charge has been that the enterprise of SSK [sociology of scientific knowledge] and the wider post-modernist fashion for textualist, historicist relativism can be seen to corrupt the standards of critical thought and honest inquiry. ... The seductive doctrines of SSK have an import going beyond their specific content. ... Latour and Woolgar describe science as the social construction of fictions. ... There is a disturbing affinity between such views and those of revisionist historians who would deny the gas chambers of Auschwitz. This kind of historical relativism serves to illustrate the consequences of the wholesale rejection of philosophical notions such as truth and reality. ... This responsibility of intellectuals to speak the truth and to expose lies is also the responsibility of teachers to instill. Above all, it is the responsibility of science teaching to convey what Bronowski called 'the habit of truth' which is central to the scientific enterprise. (Slezak, 1994, pp. 289-290)

Slezak is not alone in his distaste for extreme relativism. His views on this subject are shared by many scholars, including myself and Laudan (1990, p. x) who declares that "The displacement of the idea that facts and evidence matter by the idea that everything boils down to subjective interests and perspectives is...the most prominent and pernicious manifestation of anti-intellectualism in our times." Nevertheless, it is disturbing to see large segments of the intellectual community either approving of radical relativism, or not actively arguing against it.

In contrast with extreme interpretations of logical skepticism or cultural influence, Sections 2.44-2.45 describe a 'critical thinking' treatment of scientific method (and of its critics) that I think is more accurate and would be more beneficial for students.[30] It suggests, for example,

---

[30]. Additional elaboration by myself appears in Appendices B41-B47. And an informative discussion of controversial issues is found in Matthews (1994), which provides a good outline of the major issues: rationality, student motivation, debates among scholars, radical

distinguishing between a quest for certainty and for the 'rationally justified confidence' that is the goal of modern scientists. And to help students think more carefully, teachers can use language more carefully, with increased precision such as a distinction between theories and observation-theories, and with the use of terms (such as 'status' and 'critical realism') that allow flexibility and don't force thinking into dichotomous channels.

Above all, in an effort to act wisely, motivated by an awareness of our responsibilities as educators, we should ask "What effects will our educational policies have on students and society?" This is important for all students, but especially for younger students who are not as well equipped to intellectually defend their own ideas. An essential ingredient in the art of teaching is to judge the intellectual sophistication of students, and then use this knowledge to adjust the demands for conceptual acquisition and critical thinking to an appropriate level.

This section began by asking "What perspective on science is most accurate, and most beneficial?" These important questions are worth asking, even though (or possibly because) there are no simple answers. Instead of seeking a definitive solution that will satisfy everyone, which is impossible, the goal of this question-asking should be to stimulate a thoughtful evaluation of the relative merits of different approaches to teaching the nature of science, and to consider how these evaluations are influenced by our individual and collective perspectives on the complex relationships between *models of science, quality of education,* and *quality of life.*

4.3: Using ISM for Teacher Education

Teacher education is an important component in the development and application of a curriculum. For example, part of the difficulty in implementing the newly developed science curricula of the 1960s was the lack of preparation by teachers, who were not able to teach the courses in the way intended by the course designers:

> Many teachers, particularly in elementary and junior high schools, but including many in high schools as well, did not have sufficient training in science and math to handle the new materials. (Jackson, 1983, p. 152)

---

constructivism, feminism, multicultural education, realism, idealism, and more.

An important part of preparation is to help teachers improve their ability to "cope with complexity" and to more effectively use instructional tools such as ISM.

Recent proposals (e.g., Matthews, 1994) have suggested that teacher education could be improved by a more effective use of insights gained from history and philosophy of science:

> A teacher ought to know more than just what he or she teaches. As educators, teachers need to know how this knowledge has come about, how its claims are justified and what its limitations are. Teachers should have a feeling for, or appreciation of, the tradition of inquiry into which they are initiating students. (Matthews, 1994, p. 213)

In the types of teacher education programs envisioned in these proposals, ISM could serve a valuable function by providing an integrated structure that would help teachers develop a deeper, more coherent understanding of the problem-solving methods used by scientists.

4.4: General Thinking Skills and a 'Wide Spiral' Curriculum

This section is inspired by the important question of how thinking in science is related to thinking in other areas of life. Section 4.41 describes a prototype model for an 'integrated design method' that is more general than an 'integrated scientific method'. Section 4.42 discusses the potential for developing a 'wide spiral curriculum' that encompasses science and other areas. And Section 4.43 concludes this dissertation with a brief summary of the potential uses of ISM in education.

4.41: A Model for an 'Integrated Design Method"

There are many similarities, and some differences, between the methods used in science and the methods used in engineering, medicine, architecture, mathematics, music, art, literature, philosophy, history, business, and law, in everyday common-sense reasoning, and in learning physical skills. One way to interpret similarities and differences is to think in terms of 'design', and then compare the goals in different areas. A useful distinction can be made between goals that result in the design of products and the design of strategies. Some typical *products* are machines (designed by engineers), repaired machines (by mechanics), legislation (by politicians), songs (by musicians), backhands (by tennis players), and theories (by scientists). *Strategies* for action can be designed for a wide variety of situations, including political, social, romantic, diplomatic, scientific, athletic, military, legal, and financial. In science and in other areas, the process of design is guided by goal-criteria that specify the characteristics of a satisfactory product or strategy, and by method-criteria for the process itself, such as limitations of time and money.

I have developed a model that describes a generalized *integrated design method* (IDM) in terms of feedback loops between the goal-criteria for a product (or strategy) and the results of goal-directed design that is done mentally (using mental experiments) or physically (using physical experiments). This IDM model (depicted in Figure 9) is similar to ISM, sharing key elements and relationships. And to build decontextualized 'concept bridges' that promote forward-reaching transfer between science and other areas of design, a common symbolism, both verbal and visual, is used in ISM and IDM.

But there are also differences. Because science is just one of many areas involved in design, ISM (for only science) is a special case of IDM (for all design), so ISM can be specialized and detailed, while IDM is more general and simple, with a greater flexibility so it can adapt to a wide range of goals for different types of design. There is also an essential difference in the goals for science and for other areas, and this affects the goal-seeking methods that are used. For example, the experimental systems designed by a scientist, whose main goal is to understand how a system works, will usually differ from those designed by an engineer whose main goal is to build an improved system; Schauble (1991) describes some ways in which these differences in goals affect the behavior of students who are solving problems that are defined (by the teacher) as having a 'science goal' or an 'engineering goal'. As usual, when comparing different areas of knowledge it is instructive to search for both similarities and differences. The similarities will call attention to the many opportunities for transfer, and the differences will help students appreciate the unique characteristics of each field.

4.42: A Wide Spiral Curriculum

ISM or IDM, individually or combined, could be used as the basis for a *wide spiral* curriculum. A coordination of learning over a range of subject areas — including all sciences and some nonscience areas — would be the 'wide' aspect, and distribution of learning over time would be, using a conventional term, the 'spiral' aspect. Distribution over time occurs in a short-term spiral when the same themes are repeated and coordinated (with respect to different types of experience, levels of sophistication, and contexts) in one course. If the learning opportunities in this course are coordinated with those in other courses a student is currently taking, and if this wide approach is continued for a long period of time, the result will be a long-term, wide spiral curriculum. A well designed 'wide spiral' would have a carefully planned sequencing and coordinating of activities within each course and between courses, in science and in other areas, to form a mutually supportive, synergistic system for learning higher-level thinking skills.

As with any new curriculum, the design of a wide spiral should be guided by practical constraints, such as accommodating a wide range of teaching styles and student learning styles (especially in the culturally diverse, decentralized system of American education) and minimizing the extra 'preparation time' for teachers by helping them quickly learn the new instruction methods. Perhaps the most difficult challenge is simply making a decision to invest the classroom time that is required to teach thinking skills, despite pressures to cover a large amount of content. These are tough problems, and many educators have been (and will be) struggling with ways to achieve reasonably satisfactory solutions. Of course, any attempt to construct a 'wide spiral' should be coordinated with the work of other educators, including those (such as Nickerson, 1988; Marzano, et al, 1988) who have made recommendations for incorporating 'thinking skills' throughout the curriculum.

Returning to a more concrete level, student experiences in a wide spiral classroom could include a variety of activities such as laboratory work, group projects, discussions and debates, pauses for metacognitive reflection, problems with vaguely defined or conflicting goal criteria, science-technology-society issues, problems where students can play the role of detectives, case studies from current events or history of science, research using computer simulations of natural systems or designed systems, and 'direct learning' by reading or listening.[31] One way to coordinate

---

[31]. Yes, *direct learning* is an activity because — when done well, with the intention to learn and to integrate new material with existing personal schemas — it must be a cognitively active, constructive process. It is also an essential part of science; content-knowledge is necessary

different areas, even within one course, would be to supplement a conventional 'STS' approach by using ISM (for science) and IDM (for technology and science) to show the characteristics of Science and Technology, and their mutual interactions with each other and with Society.

One potential approach, which might be relatively easy to implement because it blends smoothly with conventional teaching practices, involves a modification of typical labwork. In a modified lab, the traditional physical activities could be combined with the mental actions needed to pose and probe simulated research problems. This first-hand experience could be supplemented with second-hand stories of scientists and their problem-solving activities. When used for instructional design, ISM and/or IDM could stimulate and structure a search for Aesop's Activities to provide opportunities for a wider range of experience. And when used in the classroom, ISM or IDM could help students learn more from their experience by directing attention to what can be learned from each experience.

Section 4.21 discussed the valuable functions that ISM could serve by providing a structure for the integration of problem-solving activities, and a motivation for intentional learning and forward-reaching transfer. By supplementing ISM with IDM, this potential value is increased by providing a wider range of integration and motivation. For example, students who do not plan to become scientists will be more motivated to learn scientific methods — to eagerly ask themselves "What can I learn now?" — if they have a reason to intentionally learn, if they see a close connection between science (in ISM) and their chosen area (in IDM), if they see a personally meaningful reason to seek forward-reaching transfer from 'scientific skills' to general 'thinking skills' that can be used outside science.

4.43: In Praise of Variety in Education

As stated above, "perhaps the most difficult challenge is simply making a decision to invest the classroom time that is required to teach thinking skills, despite pressures to cover a large amount of content." The use of higher-level thinking skills throughout a curriculum has advantages, but in

---

(but not sufficient) for thinking creatively and critically. And most scientists learn most of what they know about science content by direct learning.

doing this there is a danger that skill-learning will become the loser in a contest of priority with content-learning. In education, and especially in science education, when there is a limited amount of time for instruction there often seems to be a conflict between the teaching of process and content. One approach, taken by many educators including Perkins & Simmons (1988), is to explain how both types of learning are closely related, and how thinking skills can promote the learning of content. Another approach is to emphasize the importance of thinking skills, to argue that these skills are worth developing for their own sake, that it is worthwhile to invest some time in specialized courses designed for the specific purpose of developing skills in productive thinking. The development of thinking skills would be my main reason for advocating the type of instruction used in the MG genetics course. Earlier, in the introduction to my "suggestions for improvement" in Section 3.52, I praised this course and its teacher, and explained why I think that every student should take this type of course at least once. But I don't think it would be beneficial, for any student, if every course was taught like the MG course. The next two paragraphs explain why.

The MG course gives students an opportunity to learn valuable skills that usually are not learned, and probably cannot be learned well, in conventional courses. But this does not mean that these "opportunities for experience" should be repeated over and over, using the same type of instruction. Often, much of what can be learned from a set of experiences is learned during the first opportunity to learn. For example, my high school civics teacher mixed conventional content-oriented teaching with debates. He was a skilled debater who took one side of an argument on Monday and convinced most of us that "his side of the issue" was correct. But on Tuesday he argued for the other side and we discovered that on the previous day he had ignored many relevant facts and logical arguments. After a semester of this we were more street-wise in logic and rhetoric, and in making policy decisions, having learned that we should get the best information that all sides of an issue claim as support; only then is it possible to reach a responsible conclusion. This experience also taught us that sometimes it is wise to change opinions when new evidence becomes available, or at least to admit that people on "the other side" have good reasons for believing as they do. These valuable lessons have remained fresh in my mind to this day, and I'm glad I took this course once. But to experience only debates, every day in every course, would not produce a well balanced education. Much of what could be learned from a debating format, I learned from this one set of experiences. Although most other courses could benefit from the use of some debating, to supplement other activities that include non-competitive discussions where "winning" is not the goal, I don't think debating should be the main format, and certainly it should not be the only format.

Similarly, instruction such as that in the MG course, based on "guided simulated research experience," can be a valuable component of a curriculum, even if (or especially if) most of the remaining instruction uses other instructional approaches.[32] For students who want to become scientists, professional success depends on a combination of science-content knowledge, science-process skills, social skills, and motivation. The MG course seems to be an excellent way to improve three of these four components, but is not an efficient way to learn content. Compared with the MG course, students could learn more genetics knowledge, more thoroughly, in nine weeks of full time, high-quality conventional classroom instruction. But the MG course is extremely valuable for learning process skills, and for motivation. It is also useful for showing students the importance of learning science content deeply, because students have seen the value of learning concepts in a way that lets them "think with their knowledge." More important, however, students can learn about the nature of science, and they can experience the joy of science, which will provide motivation for learning content in
conventional courses and (especially) from textbooks.[33]

And for students who do not plan to become scientists, the MG course is valuable because it helps them learn about the nature of science, including the experience of having fun while doing science, so they can understand why scientists enjoy it. In fact, for a nonscientist who takes only one science course, the best choice might be one based on 'simulated research' experience, such as the MG course. But a wider variety would be obtained by including several approaches — such as basic 'scientific literacy', STS (Science, Technology and Society), and simulated research — in a

---

[32]. In this case, because the skills used in the MG classroom are rarely practiced in other courses, educational value of these experiences is increased. An eclectic approach to curriculum design has many advantages. Some of these are discussed in an earlier paper (Rusbult, 1992), along with "Disincentives to Eclecticism" in the paper's appendix.

[33]. Usually, books are the source of most content knowledge. For example, I know a lot about chemistry, and almost all of it was learned by an "active" reading of textbooks.

time period longer than nine weeks. For each of these approaches, ISM and/or IDM could serve a number of important educational functions.

4.5: An Overview of "ISM in Education"

Sections 4.1-4.4 have described, very briefly, some ideas for potential educational applications of ISM. But none of these ideas has been fully developed or adequately tested, so for each idea a humble disclaimer is appropriate. For example, at this time IDM is only a prototype that can serve as a starting point for further development. Section 4.42 describes some general characteristics of a 'wide spiral' curriculum that could utilize ISM, IDM, or ISM-and-IDM, but I have offered no detailed plans for designing such a curriculum or for implementing it. And the paragraph describing "practical constraints" is, when placed in proper perspective, only a humble recognition of the many problems associated with curriculum reform, rather than a claim to have solved any of these problems. Still, I think that ISM (with or without IDM) has potential, and could make valuable contributions to education.

Many possible benefits of using ISM have been discussed in this chapter. One type of application is the use of ISM, by curriculum developers or teachers, to analyze or design instruction; in this case any educational benefit to students would be indirect, through the improved instruction that might result. Similarly, if ISM was used in teacher education programs and if this helped teachers improve their understanding of science or their ability to share this understanding with students, the benefits for students would be real yet indirect. But students would benefit directly from the classroom use of ISM if this helped students improve their skills in thinking productively and in constructing mental models of science that are more clear, complete, accurate, and integrated. Thus, whether an application of ISM directly involves curriculum developers, teachers, or students, the ultimate goal is improved education for students.

My experience with a model of 'integrated scientific method' has shown it to be a powerful tool for describing the characteristics and interrelationships of activities used in the problem-solving methods of science, and for the analysis and design of instructional methods. This model, utilized for a science course or a ‘wide spiral’ curriculum, might play a valuable role in promoting the learning, retention and transfer of thinking skills. One useful characteristic of ISM is its central location at the intersection of many disciplines and the diverse perspectives they encompass. The centrality and inclusive nature of ISM could facilitate a cooperative sharing of ideas among scholars who are involved in science, science education, and the study of science. ISM can easily connect with the large amount of thinking that has been done about the methods of science and their application to education; and the familiarity of ‘scientific method’ will make it easier to use ISM as a vehicle for communicating ideas. In summary, ISM seems to have the potential for stimulating many innovative, productive ideas. Further development of this potential, in coordination with the work of others, is likely to be rewarded in the form of education that will help students improve their skills in creative and critical thinking.