DOCUMENT RESUME

ED 305 928 IR 013 797

AUTHOR	Mandinach, Ellen B.; And Others
TITLE	The Impact of the Systems Thinking Approach on Teaching and Learning Activities.
INSTITUTION	Educational Testing Service, Princeton, N.J.
REPORT NO	ETC-TR-88-31
PUB DATE	Dec 88
NOTE	102p.
PUB TYPE	Reports - Research/Technical (143)
EDRS PRICE	MF01/PC05 Plus Postage.
DESCRIPTORS	*Cognitive Processes; *Computer Simulation; *Curriculum Development; High Schools; History Instruction; *Instructional Effectiveness; Learning Activities; Problem Solving; Psychological Studies; Science Instruction; *Systems Approach
IDENTIFIERS	*Systems Thinking and Curriculum Innovation Proj

ABSTRACT

The Systems Thinking and Curriculum Innovation (STACI) project is a multi-year research effort intended to examine the cognitive demands and consequences of learning from a systems thinking approach to instruction and from using simulation-modeling software. This study tests the potentials and effects of integrating the systems approach into science and history courses to teach content knowledge as well as general problem solving skills. Also examined is the effectiveness of STELLA, a simulation-modeling software program, as a tool by which to examine scientific and historical phenomena. The focus of the research is on the learning outcomes and modeling performance that result from introducing an instructional environment that enables students to learn from and make concrete multiple representations of dynamic phenomena. Analyses of the data are presented in 14 tables and 12 appendixes include additional data, worksheets for various tasks, and a description of ways in which the systems thinking approach was integrated into three science courses. (16 references) (Author/EW)

The Impact of the Systems Thinking Approach on Teaching and Learning Activities

December 1988

Prepared by

Ellen B. Mandinach

Margaret E. Thorpe

Colleen M. Lahart

Educational Testing Service

Princeton, New Jersey

STACI Project

<u>Group Members</u>

Nancy L Benton

Hugh F. Cline

Colleen M. Lahart

Ellen B. Mandinach

Margaret E. Thorpe

The Impact of the Systems Thinking Approach on Teaching and Learning Activities

The Systems Thinking and Curriculum Innovation (STACI) Project is a multi-year research effort intended to examine the cognitive demands and consequences of learning from a systems thinking approach to instruction and from using simulation-modeling software. The purpose of the study is to test the potentials and effects of integrating the systems approach into science and history courses to teach content knowledge as well as general problem solving skills. The study also examines the effectiveness of using STELLA, a simulation-modeling software program, as a tool by which to examine scientific and historical phenomena. The research focuses on the learning outcomes and modeling performance that result from introducing an instructional environment that enables students to learn from and make concrete multiple representations of dynamic phenomena.

General Background Information

The intent of the Systems Thinking and Curriculum Innovation (STACI) Project is to examine the teaching and learning activities that result from introducing a systems thinking approach to instruction in secondary school science and history. As defined here, the systems thinking approach consists of three individual but interdependent components. First, there is system dynamics, the theory on which the instructional perspective is based. The second component is STELLA (Structural Thinking Experimental Learning Laboratory with Animation; Richmond, 1985; Richmond & Vescuso, 1986), a software package that can be used as a tool to teach systems thinking, content knowledge, and problem solving. The third component is the Macintosh microcomputer on which the STELLA software runs. The research focuses on the curriculum development that integrates the systems approach into existing courses and the learning outcomes that result from using such an approach and software in classroom settings.

The STACI Project, which began during the 1986-1987 academic year, was a two-year research effort now concluding its final year. The study was conducted at Brattleboro Union High School (BUHS), Brattleboro, Vermont in which four teachers used systems thinking in their courses. Content areas included general physical science (GPS), biology, chemistry, physics, and an experimental history course entitled War and Revolution. The purpose of the project was to examine the extent to which students acquired higher-order cognitive skills through exposure to and interaction with curricula infused with systems thinking and subsequently generalized knowledge and skills to tasks in other substantive areas. Comparisons were drawn between traditionally taught courses and those that used the systems approach. The research enabled the examination of skill and knowledge transfer across content areas as students were exposed to courses that used the systems approach.

Three ancillary studies were conducted in conjunction with a main classroom study and reported elsewhere. The first substudy (Mandinach, 1988b) focused on a select group of students who received extensive exposure to the systems approach in the War and Revolution seminar. These students were studied in an intensive case study format. The objective of this study was to collect indepth information about the students' thought processes, performance patterns, knowledge, and general problem solving skills. The second substudy (Cline, 1988) examined the organizational impact of the introduction and implementation of systems thinking in the high school. The objective was to analyze changes that occur in the structure and functioning of the school as a result of the curriculum innovation. The third substudy (Mandinach, 1988a) was a clinical cases study and cognitive analysis of students' performance and acquisition of higher-order thinking skills on a subsample of chemistry and physics students.

The purpose of this document is to report on the curriculum development, teaching activities, and learning outcomes that resulted in two years from the integration of the systems approach into several high school science and history courses. We provide descriptions of systems thinking, the site, the design, data collection, instrumentation, curricula (Appendix M), and cognitive outcomes from the main classroom study.

The Systems Thinking Approach

Systems Thinking

Systems thinking is a scientific analysis technique that provides a means by which to understand the behavior of complex phenomena over time. In recent years appreciation has developed particularly for the heuristic value of systems thinking. The creation and manipulation of models is increasingly recognized as a potentially powerful teaching technique. Based on the concept of change, system dynamics uses simulations and computer-based mathematical models to represent complex relationships among variables in the environment (Forrester, 1968). It is possible to understand the rule-like behavior of systems by constructing models of variables and their interactions, and examining the cause-and-effect relationships among the variables. The notion of a system is based on: (a) variables that characterize a system and change over time; (b) relationships among variables are interconnected by cause-and-effect feedback loops; and (c) the status of one or more variables subsequently affects the status of other variables.

Simulation models, simplified representations of systems over hypothetical time, are used to examine the structure of systems. Using simulations, characteristics of selected variables can be altered and their effects on other variables and the entire system assessed. To build a simulation, it is necessary to understand the major variables that comprise the system. These variables are used to form dynamic feedback systems, expressed in simultaneous equations. Over time, variables change and subsequently cause other variables and their interactions to change. Thus, systems focuses on the connections among the elements of the system and provides a means to understand how the elements contribute to the whole (Roberts, Andersen, Deal, Garet, & Shaffer, 1983).

STELLA and the Macintosh

The concepts that underlie the field of system dynamics form the basis for much of the simulation software that currently is used in educational settings. Until recently, the instructional use of systems thinking was constrained to environments that had powerful mainframe computers. The advent of a new software product, STELLA, has made it possible to operationalize these concepts on a microcomputer in a user-friendly environment (n.b., microdynamo has been available but language constraints make it particularly cumbersome and difficult to use). STELLA capitalizes on the graphics and icon technology of the Macintosh microcomputer (several windows--structural diagrams, equations, graph pads, tables--are available to the learner), thereby enabling individuals not versed in the intricacies of mathematical modeling to create their own systems. By minimizing the mathematical and technical skills needed to construct models, STELLA facilitates the creation and manipulation of complex models of system phenomena.

STELLA facilitates student introductions to analytic problem solving perspectives inherent in systems thinking through an iterative process of simulation model construction. Model-building requires learners to formulate, test, and revise hypotheses about relations within the dynamic systems. Modeling, as conceptualized in STELLA, is a three-step process. First, learners use STELLA's "tool kit" to create diagrams representative of the systems to be modeled. These diagrams are based on relational assumptions and logic hypothesized by the users. As learners create structural diagrams, STELLA translates them into sets of equations that represent the systems. Second, learners then formally specify the logic that connects the parts of the systems. STELLA's graphics capabilities facilitate this logical translation by providing visual maps of the connections among the components of the systems. The simultaneous equations are based on the mathematical assumptions and values learners supply to define the variables and the relationships among them. Finally, STELLA dynamically runs the systems as simulations over hypothetical time, given the logical assumptions provided by the learners. Results of the simulation stimulate revisions, thus creating an iterative sequence of formulate-test-revise steps that lead to the development of dynamic models of phenomena over time.

Site Description: Brattleboro Union High School

BUHS serves a rural five-town district in southeastern Vermont whose population is approximately 20,000. The school has roughly 1,600 students and a faculty of 80 teachers. Since 1985, BUHS has been the site of a number of systems thinking activities, all with the purpose of introducing students, educators, and the public to the principles that underlie the field (Mandinach & Thorpe, 1987).

Four teachers formed the systems group at BUHS. All were trained to use the systems approach and integrated this perspective into their courses. One course, entitled War and Revolution, was heavily infused with systems thinking and the use of STELLA. Systems thinking formed the basis for this course. In contrast, an integrative approach was used in the science courses. The approach was integrative in that the teachers identified concepts within their curricula that could be enhanced by the use of systems principles. Rather than teach particular concepts as they had in the past, the systems teachers explicitly emphasized the systemic nature of the topics, noting such ideas as causality, feedback, variation, and interaction. The courses covered the same body of knowledge taught in the traditional curriculum, but specific concepts were discussed from a systems perspective. These courses will be described in greater detail in the curriculum section.

Design and Data Collection

Design

In the 1986-1987 academic year, systems thinking was integrated into three GPS, four biology, and three chemistry classes (Table 1). An equivalent number of traditional (control) courses were taught concurrently by other members of the faculty. In the project's second year, the systems approach was used in two GPS, four biology, three chemistry, and three physics classes. The traditional treatment contained one additional class per subject, but no physics classes.

Data Collection: Instrumentation

Several types of instruments were used to assess ability, content-specific knowledge, systems thinking, and higher-order thinking skills in various stages of the research. These instruments included pretest, in-class topic, and posttest measures. General information about the instruments can be found in Appendix A.

Pretests were used to assess subjects' ability and content-specific knowledge. BUHS supplied the students' most recent standardized achievement test scores. The California Achievement Tests served as rough estimates of general ability. ETS also administered a small battery of tests, including the Advanced Progressive Matrices (Raven, 1958, 1962), to provide another index of general ability. Other measures related to skills hypothesized as important concepts underlying systems thinking were given. These included inductive and deductive reasoning, figural analogies, and understanding relationships and were parallel forms to the tests administered in the project's first year.

Modified versions of previous final examinations were given to both the systems and traditional classes. The GPS, biology, chemistry, and physics teachers took last year's tests, identified critical, yet basic concepts, and gave the shortened versions to their classes early in the academic year. These tests served as baseline assessments of content knowledge in the subject areas.

Teachers gave content tests and exercises in their courses during the year. Teachers attempted to make these activities comparable for the systems and control classes in their subject coverage. In GPS, the systems and traditional teacher collaborated on a test of Speed and Motion. In biology, the systems teacher and the second traditional teacher collaborated on transport and immunology tests. The common chemistry test was on rates of reaction. Each systems teacher also prepared an open-ended problem that related to science content covered by systems modules. These problems were given to both systems and traditional classes.

The teachers also asked teachers to prepare and administer common final examinations to their classes so that we would be able to compare differences in content knowledge that resulted from using the systems approach. Posttests were given by all the systems teachers and the traditional GPS teacher. Unfortunately, the traditional chemistry and biology teachers failed to collect these critical data. However, from the end-of-year test given, we were able to match in content coverage, enough biology items across teachers from which to make comparisons.

ETS developed a 68-item instrument that was used to assess knowledge of systems thinking and STELLA. The instrument contained items of increasing difficulty that measured a broad spectrum of skills along a continuum ranging from elementary concepts to complex modeling skills. The 1987-1988 version of the Systems Thinking Instrument (STI) was a revision of the test piloted last year. It was based on input from the BUHS teachers and systems experts and through rational task analyses. Measures of systems thinking focused on concepts such as knowledge of graphing, variation and variables, causation and causality, feedback, looping constructs, modeling, and STELLA. This test was administered at the end of the year to only the systems thinking classes.

Data Collection: Observations and Interviews

Classroom observations were conducted during a number of site visits during the two-year project. Three project members observed both systems and traditional classes to obtain information about course content, structure, and classroom procedures. Systems classes were observed when systems modules as well as traditional materials were presented. Observations were scheduled when similar topics were covered to see how the systems and traditional teachers differed in their approach to the highlighted concepts. For example, we observed how the chemistry teachers presented the topic of reaction rates, noting differences in emphases, presentation, and other areas due to the use of systems concepts.

Interviews with systems and traditional teachers were conducted to obtain additional information about the classes. It was critical to gather information from the systems teachers concerning the issues they confronted during the implementation of the curriculum innovation. The interviews also provided an opportunity to probe teachers about their perceptions of the systems thinking modules, implementation difficulties, and other issues related to the effects of the curriculum innovation on their teaching activities.

To examine variation in content emphasis, all science teachers were asked to provide information on the amount of coverage of different curriculum topics. The systems teachers also were asked to indicate the time devoted to instruction in systems thinking and to which topics the approach was applied.

Results

Discussions of the results will focus on four types of data. First, ability test performance is reported. These data include the California Achievement Test (CAT) and the Advanced Progressive Matrices (Raven, 1958, 1962) which served as measures of general ability. Other tests from the reference battery served as measures of skills hypothesized to be related to systems thinking. Third, performance on the STI is reported. A complete description of the rationale for and content of this test is discussed. Finally, performance on content tests in GPS, biology, chemistry, and physics is documented. A discussion of the relationship of these tests to the systems thinking curricula also is provided.

The various analyses reported here require examination of different parts of the sample. Consequently, sample sizes differ in accord with the measures used in the analyses. Two caveats should be noted. First, because of concurrent enrollment in more than one systems class, special treatment of the data was necessary to account for overlapping courses. Thus, scores are sometimes reported for the science courses without physics. Second, it should be noted that students in the War and Revolution seminar are included in some of the analyses to provide a comparison group (particularly for the STI). However, they are not included in the content test analyses, which focus only on the science classes. A detailed description of performance in the seminar can be found elsewhere (Mandinach, 1988b).

Ability and Achievement Measures

Two measures of ability were used to gauge students' general intellective functioning. First, students' most recent standardized achievement test scores were used to assess general crystallized ability (Cattell, 1971). Crystallized ability refers to previously constructed assemblies of performance processes retrieved as a system and applied anew in familiar instructional situations (Snow, 1980, 1982). This construct reflects long-term accumulation and organization of knowledge and skills. The reading, language, and mathematics subscales and total test score from the CAT were extracted from students' school records. To enhance interpretability, national percentiles are reported here.

Second, Parts I and II of the Advanced Progressive Matrices (APM) were used to assess general fluid ability. Fluid ability refers to new assemblies adapted to new and unfamiliar situations and novel problems by reassembly of available performance processes (Snow, 1980, 1982). For the 1987-1988 administration of the matrices, the six even items from Part I and the first nine even items from Part II were given to the students as part of the cognitive reference battery. The reliability for the instrument, using the Spearman-Brown formula for unequal lengths, was calculated at $\underline{r} = .90$.

On average, students in science ($\underline{M} = 12.11$, S.D. = 2.88) and physics ($\underline{M} = 13.06$, S.D. = 2.22) courses performed well on the APM (Table 2 and Appendix B). Students seemingly performed better in the more advanced courses with War and Revolution students scoring the highest ($\underline{M} = 14.70$) and those in GPS the lowest ($\underline{M} = 9.92$). Biology ($\underline{M} = 12.12$) and chemistry ($\underline{M} = 12.66$) students scored between the extremes. Furthermore, students in the traditional ($\underline{M} = 12.15$) and systems ($\underline{M} = 12.08$) classes performed comparably.

Data from the CAT were available for most of the BUHS students. BUHS science students averaged in the 73.31 percentile in reading, 68.68 in language, 73.75 in mathematics, and 72.66 on the total test. Appendix C presents the results of the achievement tests for the systems thinking and traditional classes in biology, chemistry, physics, and War and Revolution. GPS students scored lowest on the CAT. The data also indicated that the students in War and Revolution were outliers, compared to others in the sample as well as nationally.

Statistical analyses were performed to determine if there were pre-existing differences in achievement test scores between the traditional and systems classes within subjects. Differences were found between treatments for GPS on the CAT scales. The traditional classes were more able than the systems classes (reading $\underline{F}(1, 57) = 2.68$, ns; language, $\underline{F}(1, 57) = 8.51$, $\underline{p} < .01$;

mathematics, $F(1, 57) = 9.22$, $p < .01$; and total test $F(1, 57) = 8.15$, $p < .01$). A difference on the math subscale was noted in the biology classes, $F(2, 147) = 4.28$, $p < .05$. This difference occurred as the systems classes performed better than those taught by traditional teacher 1 ($T(92.8) = 2.86$, $p < .05$). No differences were found between systems and traditional chemistry classes.

Reference Battery

Four other tests were administered with the APM in a reference battery. The intent was to assess students' performance on skills related to systems thinking. The four tests were the Figural Analogies (FA), Diagramming Relationships (DR) and Letter Sets (LS) (Ekstrom, French, and Harman, 1976), and Deductive Reasoning (DRT). All but the DRT were parallel forms of the versions administered in the previous academic year. Results between the Year 1 and Year 2 tests are presented in Table 3. Positive correlations were found between the two administrations. The magnitude of these relationships were modest, but significant.

Figural Analogies. FA (Appendix D) consisted of 10 items presented in a four-alternative, multiple-choice format. A figural analogy item consisted of a typical analogy item (A:B::C:D), but used geometric figures, such as triangles, squares, circles, etc., rather than words. Mean performance on the items ranged from .20 to .94, with a split-half reliability of $r = .54$. One item appeared to present particular difficulty (see Appendix E).

Performance differences were noted on the FA. The systems ($M = 7.06$) and traditional ($M = 7.19$) groups performed comparably.

Students in physics (M = 7.42) scored the highest, followed by chemistry (M = 7.41), biology (M = 7.08), then GPS (M = 5.70).

Diagramming Relationships. DR (Appendix F) consisted of 15 items presented in a five-alternative, multiple-choice format. The purpose of the test was to see if students were able to discern and diagram relationships among groups of things. For example, a sample problem asks students to determine the relationships among birds, pets, and trees. The correct response would indicate that trees are neither pets nor birds, but some birds are pets. Thus, there should be an intersection between pets and birds, but not one between trees and pets or birds.

Performance on the items ranged from .34 to .93, with a mean score of 9.97. There were 10 items on which the p values were under .70 (Appendix G). The task was difficult and unlike any the students had encountered previously. The split-half reliability (Spearman-Brown for unequal parts) was r = .82.

A similar pattern of performance differences was noted on the DR. The systems (M = 9.89) and traditional (M = 9.90) groups performed comparably. Students in physics (M = 11.60) scored the highest, followed by chemistry (M = 9.88), biology (M = 9.85), then GPS (M = 7.72).

Letter Sets. LS (Appendix H) also consisted of 15 items presented in a five-alternative, multiple-choice format. Each item contained five sets of of letters in groups of four. A common rule linked four of the five letter sets. LS required identification of a rule that made four sets of letters similar in some way, and one set different from the others. For example, NOPQ DEFL ABCD HIJK

UVWX, is one of the sample items. The rule here is four letters in alphabetical sequence. The second set violates that rule, and thus is the correct response.

Performance on the items ranged from .08 to .92, with an average total score of 10.21. Performance on this task was more consistent. However, the p values indicate that students may have had difficulty finishing the task within the time constraints. Performance on Items 28, 29, and 30 reflect this possibility (see Appendix I). The split-half reliability (Spearman-Brown for unequal parts) was r = .86.

Again, a similar pattern of performance differences was noted on the IS. The systems (M = 10.21) and traditional (M = 10.22) groups performed comparably. Students in phys' (M = 10.77) and chemistry (M = 10.68) scored the highest, followed by biology (M = 9.94) and GPS (M = 8.67).

Deductive Reasoning Test. The final test, DRT (Appendix J), required students to deduce the relationships among information given in a verbal problem. The maximum score was 15. Students averaged 6.96 points. The systems (M = 7.42) students performed better than those in the traditional (M = 6.39) classes. Chemistry (M = 8.77) and physics (M = 8.56) students scored the highest, followed by biology (M = 5.54) and GPS (M = 3.64). The systems chemistry classes (M = 10.45) performed well above all other classes.

Relationship Among the Ability and Reference Battery Measures

Table 4 presents the intercorrelations among the CAT, APM, and the reference battery measures. The correlations among the reference battery measures were found to be moderate, but significant. Similar correlations were found between the reference battery and the achievement test scores. CAT total score and APM, the two primary indices of general ability, showed a moderate correlation of $\underline{r}(341) = .46$. Intercorrelations between the subscales of the CAT were extremely strong.

Systems Thinking Instrument (STI)

This section focuses on the primary instrument by which knowledge acquisition in systems thinking was assessed. The rationale for and content of the instrument will be described first. Results then will be presented, with comparisons made with the first administration.

Test characteristics. The STI (Appendix K) consisted of 68 items that were intended to assess a range of skills thought to underlie systems thinking. A rational task analysis, in conjunction with revisions from last year (see Mandinach & Thorpe, 1988), yielded eight skills which then were made into subscales of varying lengths. Table 5 presents the alpha reliabilities for those scales. The 68 items on the total test yielded an alpha of .95, indicating that the test was extremely consistent across items and subscales.

Knowledge of basic graphing concepts (e.g., labeling and scaling axes, coordinates) comprised the first subscale. The five graphing items yielded and alpha of .45. This was the least internally consistent scale. Two other scales focused on graphing skills. A first required interpretation of graphs. Students were asked to interpret a graph and provide a verbal description. The seven items on this scale yielded a reliability of .60. The second subscale required translation. That is, students were asked to take a verbal description of a problem and translate it into a graphical representation. There were six such items ($\alpha = .62$).

One other scale focused on mathematical skills that relate to systems. This scale focused on students' understanding of functions. Students were asked to write functions, find slopes, and graph functions. The three items yielded an alpha of .84.

Two scales were designed to assess knowledge of concepts critical in systems thinking. The first focused on causality ($\alpha = .94$). Here students were asked to complete a causal relationship and a causal diagram. Defining causality leads directly into the concept of looping. Four items ($\alpha = .82$) required that students either interpret or construct a causal loop diagram.

The final two scales measured skills and knowledge unique to systems, STELLA, and modeling. The first scale consisted of simple identification items. One set of identifications required students to determine if a variable was a stock (level) or a flow (rate), then define its unit of measure. A second set of items asked students to identify parts of a structural diagram (e.g., stock, connector, flow, converter). The 33 items yielded an alpha of .95. The final scale focused on the construction and interpretation of systems models of varying complexity. There were seven items in the scale which yielded an alpha of .79. At one end of the continuum, students were asked to take a simple model and identify how certain variables affect other variables. At the other end of the continuum, students were asked to take a verbal description of a problem, construct a model, and then interpret that model.

There was a rationale for constructing such a diverse test. First, we wanted to examine how the items performed and how they related to what the students had encountered in class. We needed to construct an instrument that would be both reliable and valid with respect to their classroom experiences with systems thinking. Performance on such an test would inform ETS and the BUHS teachers in terms of students' learning outcomes and cognitive processing.

Second, and perhaps most importantly, we wanted the STI to be reflective of and sensitive to differences in the systems curricula across the subject areas. As documented in a previous report (Mandinach & Thorpe, 1987), each of the teachers introduced systems concepts to their classes differently. That is, they focused on particular concepts of varying difficulty and integrated them into their courses with different degrees of intensity. Thus, the instrument should yield ranges of performance within the courses around concepts that were either stressed or briefly described. The STI was constructed to provide information about these ranges that would be reflective of the curricular differences.

Interrelationships among the subscales. Table 6 presents the correlations among the STI's scales. All correlations were found to be significant at the $p < .01$ level. The graphing scale showed moderate correlations with all other scales, whereas interpretation and translation of graphs were most strongly related to other scales. In particular, interpretation and loops ($r(212) = .48$), translation and identifications ($r(212) = .50$), and loops and identifications ($\underline{r}(212) = .46$) were most highly related. Understanding the basic concepts of systems thinking (the Systems Identification scale) was strongly related to performance on the systems thinking scale ($\underline{r}(212) = .74$).

Test performance. Item-level performance data are presented in Appendix L, which is is broken down by subject area in order to examine the progression of skills and knowledge across courses. Table 7 presents the descriptive statistics for the subscales. These data are further broken down by course in Tables 8 and 9. Whereas data in Tables 7 and 9 are presented as raw scores, percent correct is reported in Table 8 to serve as a comparison between the test's two administrations. Data discussed in the text also are reported in terms of percent correct to aid interpretability.

On all three of the graphing subscales (Graphing, Graphing Interpretation, and Graphing Translation) performance increased in the more advanced courses. Students in GPS scored the lowest, followed by biology and chemistry. The physics and War and Revolution students performed comparably. Interesting to note are the differences between Years 1 and 2. Whereas extremely high scores were found in Year 1, performance in Year 2 was substantially lower (Mandinach, 1988b). Chemistry performance was consistent across the administrations, whereas substantial improvements were noted for the GPS and biology classes.

Indepth analyses indicated particular problem areas experienced by the students. Consistent with last year, the GPS students did not understand the (x, y) convention of defining points within a coordinate system. This deficit led to poor performance on items that required graphing points and defining axes. It also led to confusions when they were asked to translate or interpret graphical problems.

GPS students continued to show no understanding of functions. They were unable to graph functions, determine slopes, or write a function when given a graphical representation. The GPS students achieved only 4 percent correct on this subscale, in contrast to 26 percent for biology, 42 percent for chemistry, 64 percent for physics, and 82 percent for War and Revolution. Functions were a difficult concept to understand and apply in graphic form.

Interesting performance patterns emerged on the two scales related to variation and causality (Causality and Loops). For both scales, the War and Revolution students achieved the highest scores, followed by those in biology, physics, chemistry, then GPS. The fact that the biology students scored higher than the other science classes on these scales is reflective of the instructional emphases stressed by the biology teacher. The items in the Causality scale required students to complete a causal statement (e.g., Amount of studying causes ______.), then make the causal loop. Last year, GPS students completed the statements, but had particular problems with the simple causal loops, obtaining only 31 percent correct. This is in contrast to 58 percent for biology, 65 percent for chemistry, and 98 percent for War and Revolution. Performance in the second year was substantially better, with GPS achieving 72 percent correct, 92 percent for biology, 86 for chemistry, 91 for physics, and 100 for War and Revolution.

Completion of more complex causal loops were required in the Loops scale. Performance was lower but analogous to the Causality scale ($\underline{M}_{GPS} = .39$; $\underline{M}_{Biology} = .73$; $\underline{M}_{Chemistry} = .60$; $\underline{M}_{Physics} = .67$; and $\underline{M}_{W\&R} = .74$). The GPS students experienced particular difficulty in translating a simple verbal description into a loop diagram. The GPS teacher admittedly devoted a limited amount of time to loops. Thus, we interpret these results to reflect the small amount of exposure the GPS students were given to loops. It is noteworthy that on this more complex task, biology students (sophomores) did as well as War and Revolution students (seniors of exceptional ability).

On the more elementary scale that tested basic knowledge of systems thinking and STELLA, the Systems Identifications scale, there were several notable trends. First, the War and Revolution class ($\underline{M} = .81$) achieved the highest score, yet at a substantial decrement from last year. These students were not able to identify inflows, outflows, or stocks, the basic elements of a systems model. In contrast, such identifications were routine for Year 1 seminar students. These differences reflect the teacher's implementation of systems in the two years (Mandinach, 1988b).

GPS students ($\underline{M} = .28$) also exhibited weaker performance in Year 2 ($\underline{M} = .42$). In contrast to the War and Revolution seminar, GPS students identified parts of a model, but had trouble with units of measure. Performance was equivalent for biology and chemistry ($\underline{M} = .63$) and slightly higher for physics ($\underline{M} = .72$). Both chemistry and biology performance improved from Year 1 to 2. Results indicated that the biology, chemistry, and physics students performed better on the structural modeling identifications and were less adept at the units of measure and variable items.

A similar performance pattern was found for the Systems Thinking scale, where the focus was on model interpretation and modeling. Given that the GPS students had little exposure to these complex concepts, they achieved only 26 percent correct. They at least made some attempt to interpret the models given to them, but had slightly more difficulty constructing their own models. The biology ($\underline{M}$ = .43), chemistry ($\underline{M}$ = .42), and physics ($\underline{M}$ = .46) classes again performed similarly, whereas the War and Revolution seminar performed only slightly better ($\underline{M}$ = .55). Improved performance was noted in GPS, biology, and chemistry. However, a substantial decrement was observed in War and Revolution. In Year 1, seminar students achieved 86 percent correct, indicating their ability to construct complex structural models. Year 2's seminar students, despite the extensive exposure to systems thinking, were unable to evidence modeling skills. Again, this is reflective of changes in the course (Mandinach, 1988b).

The contrast between Year 1 and Year 2 performance is striking (Mandinach & Thorpe, 1988). Last year, every item on the subscale exhibited the same performance trend across courses (Appendix G). War and Revolution either achieved or approached a perfect score. These scores were well above the science classes. GPS students achieved the lowest scores, whereas the biology and chemistry performed similarly. Year 2 performance was substantially different. GPS again scored the lowest on all items. However, War and Revolution performed poorly in comparison with the first class, and only slightly better than Year 2's science classes. More distressing is the relative performance on Item 28B, the complex modeling problem. Not only did the physics students score better than those in the seminar, but chemistry students performed extremely poorly. It is clear that whereas the physics and biology teachers gave attention to modeling concepts, such emphasis was lacking in War and Revolution and chemistry.

Content Tests

A number of content tests were administered in the science classes to examine the impact of systems thinking as well as to assess typical instruction.

GPS. In GPS, both the traditional and systems teachers collaborated to give 40-item pre- and posttests of knowledge of physical science. They also prepared a 30-item test on speed and motion (S&M), the topic of the primary systems module. Pre- and posttest performance were moderately correlated ($r(34) = .34$, $p < .05$). The pretest showed slightly higher correlations with S&M ($r(33) = .38$, $p < .01$) and STI ($r(31) = .36$, $p < .05$). In contrast, the posttest was highly related to performance on both S&M ($r(33) = .62$, $p < .001$) and the STI ($r(32) = .72$, $p < .001$). Students who were successful on the S&M test also performed well on the STI ($r(32) = .72$, $p < .001$).

Traditional students ($M = 16.72$) performed better than those in the systems classes ($M = 14.00$) on the pretest ($F(1,46) = 8.44$, $p < .01$). No differences were found on the posttest or S&M. However, gain score from pre- to posttesting approached significance ($F(1,46) = 3.27$, $p = .07$), indicating that the systems classes ($M = 6.56$) made substantial progress in comparison to the traditional students ($M = 4.24$).

Further analysis of performance indicated that systems and traditional students differed in their answer and error patterns on the S&M test. The traditional students excelled in their precise use of terminology. They were able to identify and define patterns of speed and motion on simple graphs. Although these students were able to recognize concepts in simple formats, they were unable to identify those same concepts in more complex problems. In contrast, the systems students were adept at decomposing parts to understand entire problems. The systems students did not concentrate on terminology. Instead, they provided descriptions of behavior over time, rather than labels for concepts.

Biology. In biology, 44-item pre- and posttests were given to assess knowledge of biological concepts. The pretest was given by all three teachers. The systems teacher administered the entire posttest, whereas parts of the final measure were given by the two traditional teachers. Thus, the samples differ, based on the content of test administration. The systems teacher collaborated with the second traditional teacher on an 11-item test to assess transport and a 31-item test to measure immunology. Fourteen selected items from the immunology test that relate to causality were separated for further examination.

Scores on the biology pretest were only moderately related to posttest performance ($r(72) = .23$, $p < .05$), indicating that those who performed well at the end of the course were not necessarily the same students who had done well earlier. The pretest was not related to performance on the immunology ($r(63) = .11$), selected immunology items ($r(63) = .09$), or transport ($r(48) = .14$) tests. In contrast, performance on the posttest was related to performance on the immunology ($r(63) = .51$, $p < .001$), selected immunology items ($r(63) = .40$, $p < .001$), and transport ($r(48) = .52$, $p < .001$) tests. Students who performed well on the transport test also did well on the immunology test ($r(45) = .58$, $p < .001$) and selected immunology items ($r(45) = .58$, $p < .001$).

Examination of the three biology teachers' classes' posttest performance indicated that the systems classes ($M = 4.96$) scored between the two traditional teachers' classes ($M_{T1} = 3.60$; $M_{T2} = 6.53$) ($F(2,155) = 32.74$, $p < .001$). Overall when the two traditional teachers' results were combined, no differences were found between performance by the systems ($M = 4.96$) and traditional classes ($M = 4.76$) ($F(1,156) = .37$, ns).

The systems and second traditional teacher gave tests of knowledge of immunology ($F(1,92) = 12.47$, $p < .001$) and selected immunology items ($F(1,92) = 6.69$, $p < .05$) on which the traditional classes performed better than the systems students. There were no differences between systems and traditional classes' performance on the transport test.

Chemistry. Systems chemistry classes received 20-item pre- and posttests, whereas the traditional classes were not given the posttest. Both classes were administered 12-item pre- and posttest on reaction rates, the topic of the primary systems module. Additional items were administered by the systems teacher to assess further the impact of the systems approach on acquisition of knowledge of reaction rates.

Performance on the chemistry pretest was moderately related to posttest scores ($\underline{r}(53) = .28$, $\underline{p} < .05$), but not related to scores on the reaction rates pre- or posttests. Similarly, students who scored higher on the reaction rates pretest did not necessarily score better on the posttest ($\underline{r}(53) = .25$, $\underline{p} < .05$). Chemistry posttest scores, however, were related to both reaction rates pretest ($\underline{r}(51) = .62$, $\underline{p} < .001$) and posttest ($\underline{r}(57) = .43$, $\underline{p} < .001$). That is, students who did better on the reaction rates tests also did better on the entire chemistry posttest.

Analyses attempted to examine the impact of course-taking sequence on performance on the reaction rates tests. Four groups of students were formed (systems chemistry, 1987-8/systems biology, 1986-7; systems chemistry, 1987-8/traditional biology, 1986-7; traditional chemistry, 1987-8/systems biology, 1986-7; and traditional chemistry, 1987-8/traditional biology, 1986-7). Differences on the pretest were not significant (Table 10). However, an ANOVA comparing chemistry treatments indicated that systems students performed better on the posttest than did those in the traditional classes ($\underline{F}(1,139) = 6.30$, $\underline{p} < .01$). Systems students also gained more points, pre-to posttesting, than did the traditional students ($\underline{F}(1,139) = 6.21$, $\underline{p} < .01$). Contrast analyses indicated that there were no differences on the posttest between systems chemistry students who had taken systems or traditional biology ($\underline{T}(102) = -.64$), nor were there differences between students who had taken systems biology students last year and were in either traditional or systems chemistry ($\underline{T}(102) = 1.62$) this year. However, students in the systems chemistry-traditional biology group performed significantly better than those in the traditional chemistry-systems biology group on the posttest ($\underline{T}(102) = 2.30$, $\underline{p} < .01$). These results may be due to the targeted nature of the curriculum modules to specific topics within chemistry and biology. Gain from pretest to posttest for these groups was not significant. Gain score for the systems-systems group was significantly better than for the traditional chemistry-systems biology students ($\underline{T}(95) = 2.33$, $\underline{p} < .05$).

Physics. The primary measures of content for physics were 39-item pre- and posttests. Students answered 12.34 of the 39 items correct on the pretest and 20.36 on the posttest, with pretest performance correlated with scores on the posttest ($\underline{r}(56) = .42$, $\underline{p} < .001$). No differences were found between students who had taken systems and traditional chemistry last year on either the physics pre- or posttests. Both groups of students gained about 7.75 points from pre- to posttesting.

Relationships Among Ability, Achievement, Content, and Systems Thinking Tests

Analyses were conducted to examine the relationships among ability, content, and systems test performance. First, the relationship between ability, as measured by the CAT, and the STI was explored. Table 11 presents the correlations between CAT and STI subscales for students in the systems classes. Performance on the mathematics scale was most strongly related to all but one STI scale. Not only was math achievement strongly related to graphing interpretation, translation, and mathematics, but also to loops, identifications, and systems thinking. Reading and language skills also were related positively to the STI scales. The graphing and causality scales were least related to CAT performance. Overall, more able students performed better on the STI. These data indicate a shift from last year's results where achievement on the CAT verbal scales was more prominent.

Second, analyses then were conducted to explore influences on content test performance in the sciences. No differences were found between GPS classes on the content posttest despite differences in ability. In comparing performance on the S&M test, controlling for ability and pretest performance, systems students performed better than traditional students. These results indicate that there were no overall content performance differences between the systems and traditional classes. However, on the topical material that was taught with the systems approach, students in the treatment condition performed better than those in the control classes.

To explore further differences in posttest performance, an ANCOVA was performed, using CAT total as a covariate. Results indicated that despite ability differences, posttest performance differed among the classes. Performance on the biology pretest was not a significant factor. Differences between the systems teacher and the first traditional teacher were not significant. However, the second traditional teacher's classes performed better than the systems classes on the biology posttest (F(1,101) = 45.04, p < .001). The ANCOVA indicated that despite ability differences, the second traditional teacher's classes performed better than the systems students.

Examination of the immunology test and selected immunology items indicated that when controlling for ability, the traditional students still performed better than the systems classes. No differences were found on the transport test when controlling for pretest and ability.

In chemistry, the systems classes performed better than the traditional students on the reaction rates posttest ($F(1,147) = 9.13$, $p < .01$). Despite ability and pretest differences, the treatment effect was sustained in the ANCOVA ($F(1,123) = 9.19$, $p < .01$). When the ANCOVA was repeated, replacing the reaction rates pretest with the chemistry pretest, the latter was found to not be a significant covariate.

A second set of analyses attempted to determine if course-taking sequence influenced performance on the reaction rates posttest. Students who had systems biology and chemistry averaged 8.43 points; systems biology-traditional chemistry 8.90 points; systems biology-traditional chemistry 7.25 points; and traditional biology and chemistry 7.74 points. These differences were not significant. However, both CAT score ($F(1,86) = 21.18$, $p < .001$) and the reaction rates pretest ($F(1,86) = 4.55$, $p < .01$) were significant covariates, with course sequence not significant.

Because there was no control class in physics, comparisons were made between students who had systems and traditional chemistry during the previous academic year. No differences in performance on the physics posttest were found nor were CAT scores or the physics pretest significant covariates.

A final set of analyses, partial correlations of ability and achievement test scores on the content tests, were performed. Both the APM and CAT total test score were used as separate partials (see Tables 12-14). Controlling for APM in the GPS classes did not substantially alter the correlations. Only when CAT score was partialled out was the relationship between performance on the STI and GPS posttest diminished (Table 12). Partialling out APM made little difference on the biology test (Table 13), whereas substantial effects were noted when CAT score was controlled. Partialling out CAT dropped most correlations substantially, particularly those that involved the STI. This result is not surprising given the high correlation between STI and CAT ($\underline{r}(41) = .58$, $\underline{p} < .001$). The correlation between STI and CAT score also was high for the chemistry classes ($\underline{r}(39) = .55$, $\underline{p} < .001$), yet controlling for CAT performance caused only two correlations (chemistry posttest-STI and chemistry posttest-reaction rates posttest) to drop substantially (Table 14). Again, controlling for APM did not alter the correlations.

Conclusions and Educational Implications

Results indicated that the BUHS teachers made substantial progress developing and implementing the systems thinking approach in their classes. Differences in approach and instructional strategies were noted from Year 1 to Year 2 (Mandinach & Thorpe, 1987). In general, changes that occurred in the science courses had positive outcomes. Differences in the War and Revolution seminar had detrimental effects on learning outcomes.

The results indicated that as the approach was integrated more effectively in Year 2, performance on systems thinking activities improved qualitatively (as reported by the teachers). The improvement was reflective of the more extensive and intensive integration of systems into the science courses. Basic graphing skills and understanding functions continued to be problematic, particularly for GPS students. However, students began to understand loops, causality, and the basics of systems thinking. Although knowledge of systems thinking concepts improved, students had difficulty modeling. Particularly distressing was the decrease in modeling performance shown by the War and Revolution students. On the positive side, the GPS and biology students improved substantially from Year 1. Chemistry performance was comparable across the two years.

Results from the content tests were mixed. In GPS, although the traditional classes performed better initially and were more able, there were no differences between groups by posttesting. There were, however, qualitative differences in the way systems and traditional students approached problems. Traditional students concentrated more on terminology, whereas systems students examined the components of behavior over time. In biology, differences among the teachers contributed to performance on the posttest, with the systems teacher's students scoring better than one traditional teacher's classes but worse than the second teacher's students. Overall, there was no posttest performance difference between the systems and traditional classes. There also were no differences on the S&M test, the measure most closely related to the systems modules. However, the second traditional teacher's classes outperformed the systems students on the immunology test and selected immunology items.

In contrast, the systems chemistry students performed better than those in the traditional classes. Furthermore, performance was dependent on whether students took systems chemistry rather than their biology treatment condition. There were no comparison groups in physics. However, students expressed that systems was a tool they could apply to solve problems in their classes (Mandinach, 1988a).

Data from the STACI Project's two years of research are somewhat mixed in their ability to provide conclusive results about the impact of the systems thinking approach on learning. Several explanations for this situation can be rendered: (a) the length of time required to develop a curriculum innovation such as the systems thinking approach and find appropriate applications for it; (b) the need to integrate the innovation into existing instruction so that students are able to make explicit ties between the innovation and traditional methods; and (c) the need for coordinated data collection that will speak directly to the research questions. The results do, however, provide valuable insights into how such a curriculum innovation takes roots, develops, and becomes a fully integrated instructional strategy.

Curriculum development made substantial progress over the project's two years. However, there is much more work that must be done to attain the desired course sequence. The teachers undertook a massive and labor-intensive project. Curriculum development is time consuming and must be done during the summer when teachers can devote blocks of uninterrupted time to work. It is impossible to learn a new body of knowledge and plan instructional materials in the 20 to 40 minute blocks of free time given to teachers during the school day. Thus it is understandable that development of the systems thinking approach has emerged slowly. The teachers worked extremely hard to identify appropriate topics to which the approach could be applied. They planned and implemented those modules in their classes. Achievement then was measured.

The research paradigm may be appropriate for many studies but does not apply in the present case. Systems thinking is an analytic technique or problem solving tool, not a content area whose acquisition can be assessed by a simple pretest/posttest design. Rather, the systems approach, as an analytic tool is expected to influence cognitive processes or the methods by which learners approach tasks. This is not to say that systems thinking will not affect the acquisition of content knowledge. The approach indirectly should influence that acquisition through the development of effective higher-order cognitive processes. Because these processes often take substantial time to emerge, it is likely that the duration of STACI provided an insufficient test of the approach's potentials. A true test of an innovation's impact must be longitudinal and make provisions for appropriate time and resources to support teachers' curriculum development efforts.

A corresponding issue is how the approach is integrated into courses, particularly sequences of courses such as in science. Students need to a sense of feel continuity and connection within and across courses. When systems thinking first was implemented, students perceived the approach to be completely independent of the science content. There was systems thinking and there was science. Systems thinking was not seen as a tool to be applied to the content to help students solve problems. To some degree, the lack of connection escalated during the second year when systems concepts were reintroduced rather than reinforced. The lesson to be learned from the BUHS experience is that explicit connections between the course content and systems thinking as an analytic tool must be built into the instruction. Furthermore, planned linkages between courses within a departmental sequence need to be developed to insure sufficient carry-over without redundancy. Students should be told explicitly about the generic use of models to solve problems and how models will be used in instruction. Furthermore, the nature of the systems activities and expectations for learning a new problem solving tool should be made explicit.

From the research perspective, it is imperative to obtain comparable data from control and experimental classes on targeted topics taught with the innovation to assess its impact. STACI lost crucial data in Year 2 that would have provided insights into the innovation's impact. However, as Cronbach (1963) noted, the power of course evaluation is not found in formal comparisons of control and treatment groups' test data where real differences get lost amid a plethora of results. Rather, "it is clearly important to appraise the student's general educational growth, which curriculum developers say is more important than mastery of the specific lessons presented" (p. 674). Cronbach recommends small, well controlled studies that enable researchers to examine alternate versions of the same course to study and can produce results with explanatory value. The three STACI Project substudies (Cline, 1988; Mandinach, 1988a, 1988b) attempt to address this need.

Another issue that emerged was the role of the teacher. The implementation of the systems thinking approach, particularly because of its theoretical and technological foundations, requires knowledge on the part of the teachers. Gaining a working knowledge of system dynamics, the STELLA software, and Macintosh is substantially different from acquiring information about another topic within content areas of expertise. Teachers need to feel comfortable with their knowledge of the approach so that they can impart that knowledge to their students. This is not to say that teachers must be experts, but rather have working experience with the approach. The STACI Project observed some reticence in the BUHS teachers. The reluctance occurred because teachers were not completely confident in their knowledge of the approach. The implementation of the approach reflected this lack of confidence.

Teacher training is a partial answer to the problem. Although all BUHS teachers received in-service training in systems thinking and one took a sabbatical to take a university-level systems course, there still were questions of knowledge, confidence, and control. Teachers were experienced but not expert. Yet they were reluctant to risk classroom control by shifting responsibility for learning to both students and the teacher.

The systems thinking approach can change the nature of the classroom. The teacher no longer is the sole expert in the classroom, with control of content, knowledge, and instructional procedures. Instead, learning becomes more interactive with shared responsibility among teacher and students. This means that teachers must take risks and alter instructional strategies to facilitate rather than simply impart learning. Not all teachers are capable of or willing to accept this evolving role.

Implementation also requires a willingness to experiment with traditional approaches. While all teachers expressed a willingness to experiment and devoted a substantial amount of time to curriculum development, there was variation among the teachers in how systems instruction was implemented. Emphasis was placed on having students think about the nature of problems by viewing interrelationships among factors. They were asked to generate explanations rather than merely identify, define, or compute solutions. Differences were evident among teachers in the use of STELLA, the type and complexity of problems presented, and the manner in which assistance was provided. The variation appeared to be attributable to a knowledge of systems theory, comfort with experimentation, and instructional style. In two cases systems instruction was implemented as a distinct change in the organization and function of the classroom. Students spent considerable time developing and experimenting online with models. These experiences were structured to engage students actively in the tasks while the teacher served as a facilitator. In another case, systems was applied more traditionally. The teacher presented systems concepts largely through class instruction in which he sought to shift students' thinking through directed questioning. A fourth teacher was severely impeded in his ability to implement systems instruction due to lack of familiarity with STELLA and the Macintosh.

In general, though, there were significant developments at BUHS as a result of the curriculum innovation. There is no doubt that the infusion of the systems thinking approach into the science and the War and Revolution seminar changed many of the teachers' instructional strategies and procedures by which to present traditional concepts. Given the ongoing nature of the curriculum development effort, it is likely that changes will continue to occur as teachers revise and implement their curriculum modules.

The results, however, do indicate that the use of the technology of the Macintosh computer and the STELLA software make accessible to students and teachers the modeling capabilities heretofore found only on powerful mainframe computers. Such modeling broadens the range of cognitive representations and instructional strategies that students and teachers can bring to bear in solving problems. The systems thinking approach, particularly with the Macintosh's graphics and mouse technology and the STELLA environment, allows students and teachers to develop and implement dynamic models of systems by using a variety of abstract representations to explore and make concrete many phenomena. It is our hope that the integration of this learning environment into curricula has the potential to produce students who have a greater capacity to understand the interrelated and complex nature of many phenomena that one encounters in daily life.

Footnote

This research was conducted under the auspices of the Educational Technology Center, Harvard Graduate School of Education and is supported by the Office of Educational Research and Improvement under contract number 400-83-0041. Any opinions, findings, and conclusions or recommendations expressed in this document are those of the authors and do not necessarily reflect the views of ETC, OERI, or ETS. The authors wish to acknowledge Hugh Cline, Nancy Benton, Lynn Patterson of ETS, Dr. Barbara Bowen of Apple Computer, Inc., Charles Butterfield, David Clarkson, Chris O'Brien, Larry Richardson, and the administration and students of Brattleboro Union High School.

References

Cattell, R. B. (1971). Abilities: Their structure, growth, and action. Boston: Houghton Mifflin.

Cline, H. F. (1988). What's happening to our schools? Princeton, N.J.: Educational Testing Service.

Cronbach, L. J. (1963). Course improvement through evaluation. Teachers College Record, 64, 672-683.

Ekstrom, R. B., French, J. W., & Harman, H. H. (1976). Kit of factor-referenced cognitive tests. Princeton, N.J.: Educational Testing Service.

Forrester, J. W. (1968). Principles of systems. Cambridge: MIT Press.

Mandinach, E. B. (1988a). Self-regulated learning substudy: Systems thinking and curriculum innovation (STACI) project. Princeton, N.J.: Educational Testing Service.

Mandinach, E. B. (1988b). The war and revolution case study: Systems thinking and curriculum innovation (STACI) project. Princeton, N.J.: Educational Testing Service.

Mandinach, E. B., & Thorpe, M. E. (1987). The systems thinking and curriculum innovation project: Technical report, part 1 (TR87-6). Cambridge: Harvard University, Educational Technology Center.

Mandinach, E. B., & Thorpe, M. E. (1988). The systems thinking and curriculum innovation project: Technical report, part 2 (TR88-12). Cambridge: Harvard University, Educational Technology Center.

Raven, J. C. (1958). Advanced progressive matrices (Set I). New York: Psychological Corporation.

Raven, J. C. (1962). Advanced progressive matrices (Set II). New York: Psychological Corporation.

Richmond, B. (1985). STELLA [Computer program]. Lyme, NH: High Performance Systems, Inc.

Richmond, B., & Vescuso, P. (1986). STELLA [Computer program] (2nd. ed.). Lyme, NH: High Performance Systems, Inc.

Roberts, N., Andersen, D., Deal, R., Garet, M., & Shaffer, W. (1983). Introduction to computer simulation: A system dynamics modeling approach. Reading, MA: Addison-Wesley.

Snow, R. E. (1980). Aptitude processes. In R. E. Snow, P-A. Federico, & W. E. Montague (Eds.), Aptitude, learning, and instruction (Vol. 1). Hillsdale, N. J.: Erlbaum.

Snow, R. E. (1982). Training of intellectual aptitude. In D. K. Detterman & R. J. Sternberg (Eds.), How and how much can intelligence be increased. Norwood, N. J.: Ablex.

Table 1
STACI Project Classes, 1986-1988

Class	1986 -1987 Systems	1986 -1987 Traditional	1987-1988 Systems	1987-1988 Traditional
GPS	3(47)	3(69)	2(41)	3(47)
Biology	4(82)	4(94)	4(80)	5(104)
Chemistry	3(63)	3(64)	3(66)	4(93)
Physics	Not included		3(65)	
Total	10(192)	10(227)	12(252)	12(244)
War & Revolution	1(8)		1(10)	

Note. Numbers indicate the number of classes (and enrollment).

Table 2

War and Revolution Students' Performance on the Psychometric Instruments: Relevant Comparisons to Systems Classes (in %)

Instrument	GPS	Biology	Chemistry	Physics	W&R
		1986-1987 Courses			
Figural Analogies	65.10	69.90	75.30		90.00
Progressive Matrices	69.87	78.73	85.53		99.07
Diag. Relationships	46.87	56.07	62.07		86.67
Letter Sets	60.60	71.73	71.73		86.67
Deductive Reasoning					
CAT Reading		70.75	84.89		95.86
CAT Language		62.00	82.98		95.00
CAT Math		68.04	81.89		98.14
CAT Total		66.60	85.11		97.57
		1987-1988 Courses			
Figural Analogies	57.03	71.16	74.92	72.20	86.00
Progressive Matrices	66.13	83.09	86.20	85.36	98.00
Diag. Relationships	51.48	67.92	62.18	76.00	87.33
Letter Sets	57.84	66.86	71.95	71.22	77.33
Deductive Reasoning	24.26	39.03	69.15	56.31	63.33
CAT Reading	55.26	69.12	76.38	85.57	94.90
CAT Language	45.30	63.90	73.92	83.88	94.80
CAT Math	54.81	73.73	76.98	84.82	95.10
CAT Total	50.74	69.03	77.14	86.78	96.60

Note. $\underline{n}_{GPS'88} = 27$. $\underline{n}_{Bio'88} = 67$. $\underline{n}_{Chem'88} = 53$. $\underline{n}_{Phys'88} = 50$. $\underline{n}_{W\&R'88} = 10$. $\underline{n}_{GPS'87} = 46$. $\underline{n}_{Bio'87} = 75$. $\underline{n}_{Chem'87} = 54$. $\underline{n}_{W\&R'87} = 7$.

Scores for the seven War and Revolution students concurrently enrolled in physics appear in W&R, not physics.

Scores for the two War and Revolution students concurrently enrolled in chemistry appear in W&R, not chemistry.

Table 3

Intercorrelations Between Reference Batteries, Year 1 and Year 2

Year 1	Year 2			
	Fig. Anal.	Matrices	Diag. Rel.	Let. Sets
Figural Analogies	.20*	.32	.30	.16*
Progressive Matrices	.30	.47	.42	.24
Diag. Relationships	.17*	.27	.43	.26
Letter Sets	.23	.22	.34	.46

Note. n = 204 - 216. * = $p < .01$.
$p < .001$ for all other correlations.
*M*FAY1 = 7.26. *M*FAY2 = 7.11.
*M*APMY1 = 12.23. *M*APMY2 = 12.23.
*M*DRY1 = 8.73. *M*DRY2 = 9.88.
*M*LSY1 = 10.93. *M*LSY2 = 10.20

Table 4

Intercorrelations Among Ability and Reference Battery Measures

	FA	APM	DR	LS	DRT	READ	LANG
FA							
APM	.40						
DR	.31	.40					
LS	.24	.33	.41				
DRT	.24	.29	.40	.29			
READ	.25	.36	.48	.26	.36		
LANG	.23	.38	.48	.36	.38	.73	
MATH	.38	.45	.52	.38	.39	.58	.66
TOTAL	.32	.46	.56	.38	.43		

Note. FA = Figural Analogies. APM = Advanced Progressive Matrices. DR = Diagramming Relationships. LS = Letter Sets. DRT = Deductive Reasoning Test. READ = CAT Reading Subscale. LANG = CAT Language Subscale. MATH = CAT Mathematics Subscale. TOTAL = CAT Total Battery Score. n = 336 to 382. $p < .001$ for all correlations.

Table 5
Reliabilities for the STI Subscales

Subscale	Items	Alpha
Graphing	5	.45
Graphing Interpretation	7	.60
Graphing Translation	6	.62
Graphing Mathematics	3	.84
Loops	4	.82
Causality	3	.94
Identifications	33	.95
Systems Thinking	7	.79
Total Test	68	.95

Note. *n* = 214.

Table 6

Intercorrelations among STI Subscales

	GR	GI	GT	GM	LOOPS	CAUS	ID
GR							
GI	.34						
GT	.31	.42					
GM	.24	.31	.42				
LOOPS	.23	.48	.34	.33			
CAUS	.31	.38	.39	.18*	.39		
ID	.26	.44	.50	.38	.46	.26	
ST	.17*	.39	.44	.30	.40	.39	.74

Note. GR =Graphing. GI = Graphing Interpretation.
GT = Graphing Translation. GM = Graphing Mathematics.
LOOPS = Loops. CAUS = Causality. ID = Identifications.
ST = Systems Thinking.
n = 214. * = $p < .01$.
$p < .001$ for all other correlations.

Table 7

Descriptive Statistics for the STI Subscales

Subscale	Maximum Score	M	S.D.
Graphing	6	4.98	1.10
Graphing Interpretation	19	11.98	4.10
Graphing Translation	17	8.98	3.60
Graphing Mathematics	5	1.82	2.04
Loops	17	10.66	4.83
Causality	9	7.82	2.34
Identifications	33	19.52	9.59
Systems Thinking	34	13.93	5.78
Total Test	140	79.70	24.43

Note. *n* = 214.

Table 8

Comparisons of Performance on the STI (in %)

Subscale	GPS	Biology	Chemistry	Physics	W&R
		1986-1987 Courses			
Graphing	57	65	80		92
Graphing Interpretation	35	52	69		98
Graphing Translation	34	47	57		91
Graphing Mathematics	8	30	53		100
Loops	27	57	62		97
Causality	31	58	65		98
Identifications	42	55	59		94
Systems Thinking	19	33	35		86
Total	32	47	56		92
		1987-1988 Courses			
Graphing	76	82	82	90	89
Graphing Interpretation	47	63	65	70	66
Graphing Translation	36	49	55	66	68
Graphing Mathematics	4	26	42	64	82
Loops	39	73	60	67	74
Causality	72	92	86	91	100
Identifications	28	63	63	72	81
Systems Thinking	26	43	42	46	55
Total	36	59	58	66	72

Note. $n_{GPS'88} = 34$. $n_{Bio'88} = 65$. $n_{Chem'88} = 60$. $n_{Phys'88} = 48$. $n_{W\&R'88} = 9$. $n_{GPS'87} = 46$. $n_{Bio'87} = 70$. $n_{Chem'87} = 56$. $n_{W\&R'87} = 7$.
Scores for the seven War and Revolution students concurrently enrolled in physics appear in both W&R and physics.
Scores for the two War and Revolution students concurrently enrolled in both chemistry appear in both W&R and chemistry.

Table 9

Course Breakdowns for the STI Subscales

	GPS		Biology		Chemistry		Physics		War & Rev.	
	M	S.D.	M	S.D.	M	S.D.	M	S.D.	M	S.D.
GR	4.56	1.31	4.91	1.06	4.93	1.19	5.39	.79	5.33	.50
GI	8.92	3.80	11.98	4.30	12.42	3.51	13.30	3.71	12.56	3.88
GT	6.15	3.87	8.29	3.20	9.42	3.03	11.18	3.03	11.56	3.05
GM	.18	.58	1.32	1.99	2.08	1.97	3.20	1.84	4.11	1.69
LP	6.65	3.74	12.45	4.12	10.27	4.48	11.37	5.17	12.67	6.18
CS	6.50	3.36	8.25	1.63	7.75	2.40	8.20	1.98	9.00	.00
ID	9.12	8.59	20.83	8.89	20.67	8.18	23.85	7.88	26.78	4.68
ST	9.03	5.30	14.72	5.50	14.27	4.81	15.57	5.89	18.67	2.60
TT	51.09	22.63	82.75	21.67	81.00	19.00	92.07	20.11	100.67	12.83

Note. GR =Graphing. GI = Graphing Interpretation.
GT = Graphing Translation. GM = Graphing Mathematics.
LP = Loops. CS = Causality. ID = Identifications.
ST = Systems Thinking. TT = Total test score.
n_{GPS} = 34. n_{Bio} = 65. n_{Chem} = 60. n_{Phy} = 54.
$n_{W\&R}$ = 9.

Table 10

Performance on the Chemistry Reaction Pates Test by Course-Taking Sequence

	M	S.D.	n
	Pretest		
Systems Chemistry	3.54	2.22	37
Systems Biology	3.05	1.93	19
Traditional Biology	4.06	2.44	18
Traditional Chemistry	3.97	1.62	62
Systems Biology	3.93	1.44	28
Traditional Biology	4.00	1.78	34
	Posttest		
Systems Chemistry	8.66	2.32	41
Systems Biology	8.43	2.69	21
Traditional Biology	8.90	1.89	20
Traditional Chemistry	7.55	2.34	65
Systems Biology	7.35	2.40	31
Traditional Biology	7.74	2.30	34
	Gain		
Systems Chemistry	5.11	3.01	37
Systems Biology	5.42	3.24	19
Traditional Biology	4.78	2.80	18
Traditional Chemistry	3.72	2.04	62
Systems Biology	3.71	2.22	28
Traditional Biology	3.74	1.91	34

Table 11

Correlations Between Achievement and Systems Performance

Subscale	Reading	Language	Mathematics	Total
Graphing	.21*	.32	.24	.29
Graphing Interpretation	.28	.38	.43	.40
Graphing Translation	.42	.47	.53	.52
Graphing Mathematics	.44	.46	.56	.53
Loops	.28	.35	.42	.37
Causality	.18*	.23	.26	.24
Identifications	.51	.53	.53	.58
Systems Thinking	.42	.43	.51	.50
Total Test	.53	.60	.65	.65

Note. * = $p < .01$. $p < .001$ for all other correlations.
$n_R = 206$. $n_L = 206$. $n_M = 205$. $n_T = 204$.

Table 12

First-order and Partial Correlations among GPS and Ability Tests

	Systems	GPS Pre	GPS Post	Speed
Systems	X	.45***	.58***	.63***
GPS Pretest	.46**	X	.43**	.42**
GPS Posttest	.66***	.45**	X	.48**
Speed	.70***	.44**	.58***	X
APM	.46**	.15	.44**	.43**
Systems	X	.68***	.48*	.61**
GPS Pretest	.70***	X	.42*	.64***
GPS Posttest	.62***	.47*	X	.46*
Speed	.62***	.65***	.47*	X
CAT	.47*	.25	.60**	.18

Note. * = $p < .05$. ** = $p < .01$. *** = $p < .001$.
$n_{APM} = 31$. $n_{CAT} = 21$.
Zero-order correlations are in the lower triangle.
Partial correlations are in the upper triangle.

Table 13

First-order and Partial Correlations among Biology and Ability Tests

	Systems	Bio Pre	Bio Post	Immun	Sel Imm	Trans
Systems	X	.24	.55***	.64***	.64***	.42**
Bio Pretest	.22	X	.19	.20	.34*	.14
Bio Posttest	.57***	.17	X	.59***	.49***	.49***
Immunology	.65***	.19	.61***	X	.87***	.56***
Selected Imm.	.65***	.32*	.51***	.88***	X	.60***
Transport	.44***	.11	.52***	.59***	.62***	X
APM	.18	-.07	.22	.21	.19	.26*
Systems	X	.16	.29*	.50***	.49***	.21
Bio Pretest	.27*	X	.06	.19	.31*	.13
Bio Posttest	.57***	.21	X	.41**	.21	.28*
Immunity	.65***	.28	.61***	X	.83***	.42**
Selected Imm.	.65***	.38**	.51***	.88***	X	.47***
Transport	.44***	.23	.53***	.57***	.62***	X
CAT	.58***	.24	.69***	.52***	.56***	.51***

Note. * = $p < .05$. ** = $p < .01$. *** = $p < .001$.
$n_{APM} = 44$. $n_{CAT} = 43$.
Zero-order correlations are in the lower triangle.
Partial correlations are in the upper triangle.

Table 14

First-order and Partial Correlations among Chemistry and Ability Tests

	Systems	Chem Pre	Chem Post	RR Pre	RR Post
Systems	X	.01	.39**	.33*	.55***
Chem Pretest	.03	X	.24	.21	-.11
Chem Posttest	.40**	.24	X	.64***	.48***
Reac. Rates Pre	.38**	.21	.64***	X	.21
Reac. Rates Post	.56***	-.10*	.49***	.23	X
APM	.35**	.04	.11	.25*	.13
Systems	X	-.08	.29*	.35*	.50***
Chem Pretest	.00	X	.12	.08	-.23
Chem Posttest	.46***	.16	X	.65***	.31*
Reac. Rates Pre	.41**	.11	.67***	X	23
Reac. Rates Post	.60***	-.16	.43**	.30*	X
CAT	.55***	.12	.44**	.23	.39**

Note. * = $p < .05$. ** = $p < .01$. *** = $p < .001$.
$n_{APM} = 47$. $n_{CAT} = 41$.
Zero-order correlations are in the lower triangle.
Partial correlations are in the upper triangle.

Appendix A
Description of Content, Reference Battery, and Systems Tests

Subject	Test	Subscore	Items	Max. Score	Teachers with data
GPS	Pretest				
	Full Year	Total	40	40	Sys., Trad.
	Posttest				
	Full Year	Total	40	40	Sys., Trad.
	Speed Motion	Total	30	63	Sys., Trad.
		Sub 1	5	5	Sys., Trad.
		Sub 2	10	10	Sys., Trad.
		Sub 3	9	26	Sys., Trad.
		Sub 4	1	2	Sys., Trad.
		Sub 5	5	20	Sys., Trad.
Biology	Pretest				
	Full Year	Total	44	44	Sys., T1, T2
	Posttest				
	Full Year	Total	44	44	Systems
		Sub 1	9	9	Sys., T1
		Sub 2	14	14	Sys., T1
		Sub 3	12	12	Sys., T2
		Sub 4	8	8	Sys., T1, T2
	Transport	Total	11	11	Sys., T2
	Immunology	Total	31	31	Sys., T2
		Selected	14	14	Sys., T2
Chemistry	Pretest				
	Full Year	Total	20	20	Sys., Trad.
	Posttest				
	Full Year	Total	20	20	Systems
	Reaction Rates				
	Pretest	Total	12	12	Sys., Trad.
	Reaction Rates				
	Posttest	Total	12	12	Sys., Trad.
	Reactions	Part 1	25	25	Systems
	Reactions	Part 2	10	10	Systems
	Reactions	Part 3	10	14	Systems
	Third Quarter	Final	25	25	Systems
Physics	Pretest				
	Full Year	Total	39	39	Systems
	Posttest				
	Full Year	Total	39	39	Systems

Appendix A, Continued

Reference Battery	Figural Analogies	10	10	All but Traditional GPS
	Advanced Progressive Matrices	15	15	
	Diagramming Relationships	15	15	
	Letter Sets	15	15	
	Deductive Reasoning	1	15	
Systems Thinking Instrument	Total	68	140	Systems GPS, Biology, Chemistry, Physics, and War and Revolution
	Graphing	5	6	
	Graphing Interpretation	7	19	
	Graphing Translation	6	17	
	Math Graphing	3	5	
	Loops	4	17	
	Causality	3	9	
	Identifications	33	33	
	Systems Thinking	7	34	

Note. Chemistry Reaction Rates Posttest is a subset of the Reactions Part 1 measure.

Appendix B
Percent Correct on the Advanced Progressive Matrices

Item	*M*	S.D.
1-2	.92	.28
1-6	.78	.41
1-10	.84	.37
2-2	.83	.38
2-6	.80	.40
2-10	.70	.46
2-14	.67	.47
2-18	.59	.49
1-4	.90	.30
1-8	.74	.44
1-12	.66	.47
2-4	.80	.40
2-8	.69	.46
2-12	.72	.45
2-16	.63	.48

Note. *n* = 401.

Appendix C
Achievement Test Scores by Teacher and Course
(In National Percentiles)

Course	Treatment	n	M	S.D.
Reading Subscale				
General Physical Science	Systems	27	55.26	20.82
General Physical Science	Traditional	32	63.84	19.39
Biology	Systems	67	69.12	21.64
Biology	Traditional	53	69.38	19.52
Biology	Traditional	33	70.30	23.07
Chemistry	Systems	53	76.66	16.61
Chemistry	Traditional	85	78.36	19.36
Physics	Systems	58	86.64	14.21
War and Revolution	Systems	10	94.90	3.57
Language Subscale				
General Physical Science	Systems	27	45.30	20.46
General Physical Science	Traditional	32	60.38	19.19
Biology	Systems	67	63.90	21.67
Biology	Traditional	52	63.85	18.20
Biology	Traditional	33	68.70	22.80
Chemistry	Systems	53	74.15	17.16
Chemistry	Traditional	85	71.44	18.61
Physics	Systems	58	84.97	15.74
War and Revolution	Systems	10	94.80	4.47
Mathematics Subscale				
General Physical Science	Systems	27	54.81	18.20
General Physical Science	Traditional	32	68.59	16.63
Biology	Systems	66	73.73	16.56
Biology	Traditional	51	63.47	21.09
Biology	Traditional	33	70.91	20.28
Chemistry	Systems	53	77.32	15.66
Chemistry	Traditional	84	78.37	15.98
Physics	Systems	58	86.12	12.32
War and Revolution	Systems	10	95.10	3.93
Total Test Score				
General Physical Science	Systems	27	50.74	18.65
General Physical Science	Traditional	32	64.03	17.09
Biology	Systems	66	69.03	20.03
Biology	Traditional	50	65.68	17.98
Biology	Traditional	32	71.34	22.26
Chemistry	Systems	52	77.46	15.11
Chemistry	Traditional	83	77.02	17.05
Physics	Systems	58	87.95	12.78
War and Revolution	Systems	10	96.60	1.96

Note. The two listings for traditional biology reflect the fact that there were two teachers in the control treatment.

Appendix D

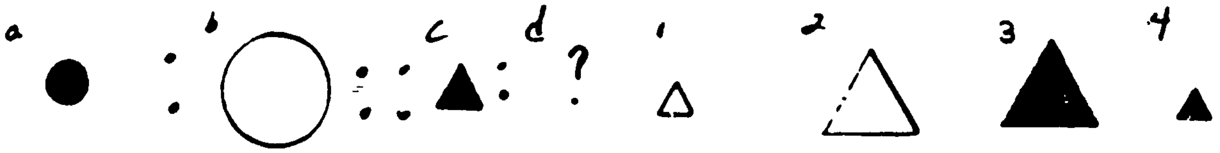
NAME______ __________ CLASS______________

FIGURAL ANALOGIES

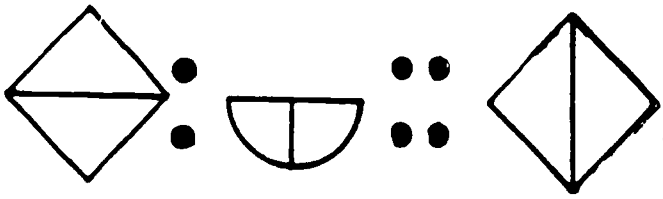
Instructions

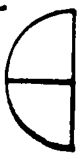
Here are several analogies in the form of figures rather than words. The format is the same. There is some relation between forms A and B with C and one of the other 4 options. Your task is to determine which of the four figures to the right completes the form, A: B :: C: D. Look at the example below The A form is a small, solid colored circle. B is a large empty circle. C is a small solid colored triangle. Which of the four figures completes the pattern? A large, empty triangle or 2 is the correct answer. Work through the items. Circle the answer that best completes the pattern.

Example:

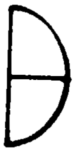


1.



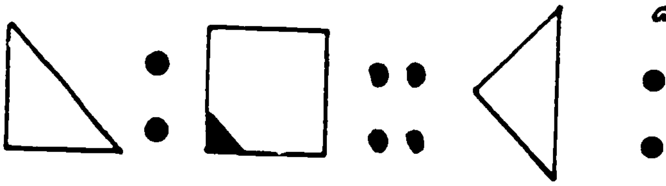
a.


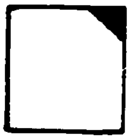
b.

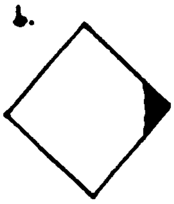

c.


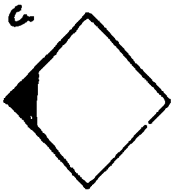
d.

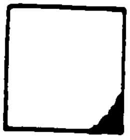

2.



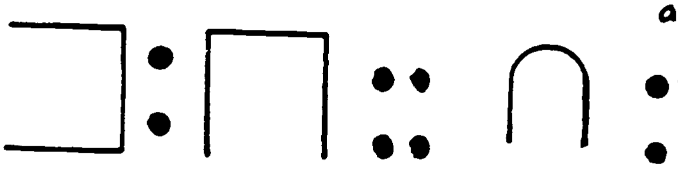
a.


b.


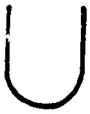
c.


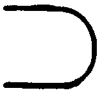
d.


3.



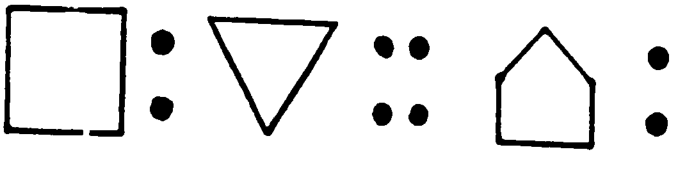
a

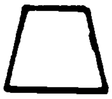

b.


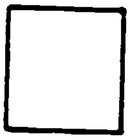
c.


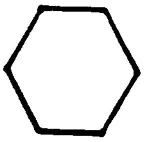
d.

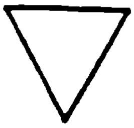

4.



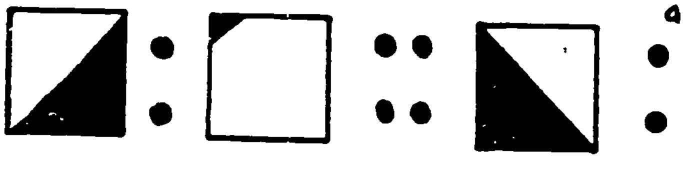
a.


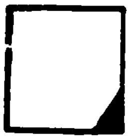
b.


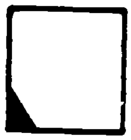
c.


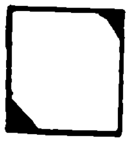
d.


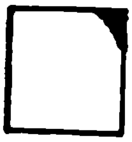
5.



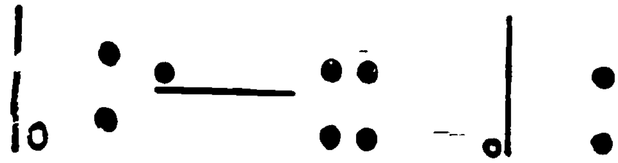
a.


b.


c.


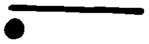
d.


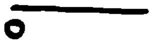
6.



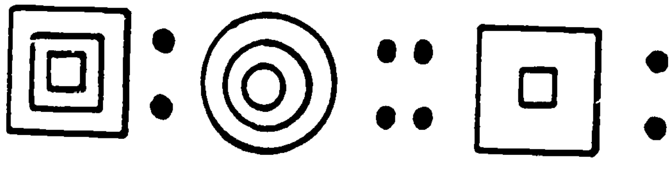
a.

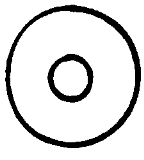

b.

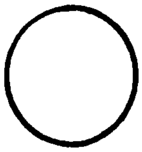

c.


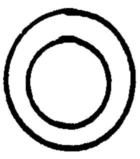
d.


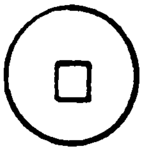
7.



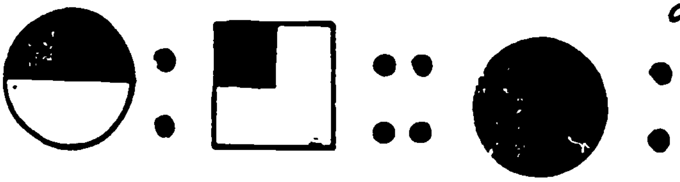
a.


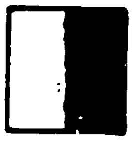
b.


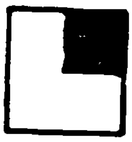
c.


d.


8.



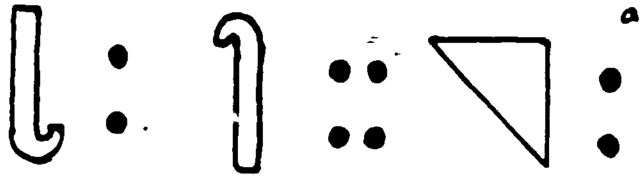
a.


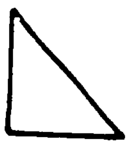
b.


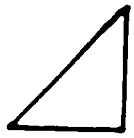
c.

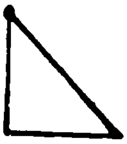

d.

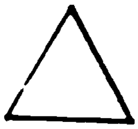

9.



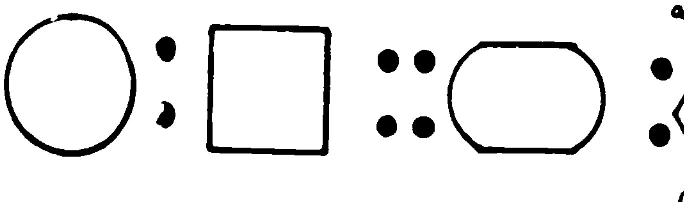
a.


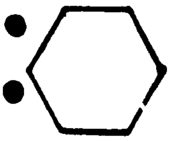
b.


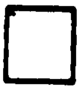
c.


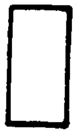
d.


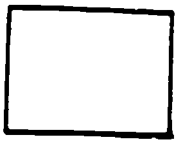
10.



a.


b.


c.


d.


Appendix E
Percent Correct on the Figural Analogies Test

Item	*M*	*S.D.*
1	.84	.37
3	.47	.50
5	.83	.37
7	.94	.23
9	.92	.28
2	.90	.30
4	.20	.40
6	.57	.50
8	.68	.46
10	.69	.46

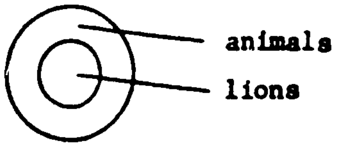
Note. *n* = 401.

Appendix F

Name ______________________________

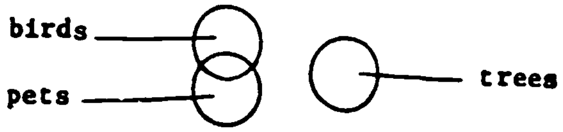
DIAGRAMMING RELATIONSHIPS -- RL-2

Sometimes the relationships among groups of things are best explained by diagrams that consist of overlapping circles. For example, if certain specific things, let's say lions, all belong to one larger class of things, let's say animals, you could diagram the situation as follows:



In these diagrams we do not care about the relative sizes of any of the circles. That is, we are not suggesting here that a relatively large proportion of animals are lions, but we are indicating that <u>all</u> lions are animals. That is why the circle representing lions is drawn entirely within the circle that represents animals.

Now take the relationships among three groups of different things: birds, pets, and trees. These should be diagrammed as follows:



This diagram shows that <u>no</u> trees are <u>either</u> pets or birds, but <u>some</u> birds are pets and <u>some</u> pets are birds.

Each item in this test names three groups of things. You are to choose from the lettered diagrams at the top of the test pages the one diagram that shows the correct relationships among the three groups of things in each item. Mark the letter of the diagram that you select.

Now try these sample items:


A


B


C


D


E

1. Animals, cats dogs

   A B C D E

2. Desks, furniture, pencils

   A B C D E

You should have marked A for 1. and E for 2.

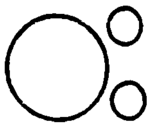
Your score on this test will be the number of correct choices minus a fraction of the number of incorrect choices. Therefore, it will <u>not</u> be to your advantage to guess, unless you have at least some idea that will help you make a correct choice.

There are two parts to this test. Each part has one page with 15 items. You will have <u>4 minutes</u> to complete each part. When you have finished Part 1, STOP. Please do not go on to Part 2 until asked to do so.

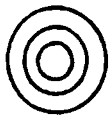
DO NOT TURN THIS PAGE UNTIL ASKED TO DO SO

Part 2 (4 minutes)

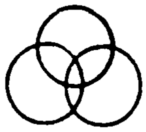
Mark the letter of the diagram that represents the relationships among the three groups in each item:



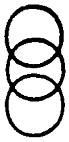
A



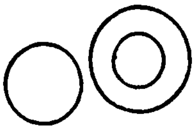
B



C



D



E

16. Shirts, things made of cotton, blue things

A B C D E

17. Pines, trees, stones

A B C D E

18. Gems, diamonds, hard things

A B C D E

19. Trousers, shirts, hats

A B C D E

20. Chairs, liquids, fishing rods

A B C D E

21. Dogs, bears, white animals

A B C D E

22. Things painted red, cars, bicycles

A B C D E

23. Storms, bad weather conditions, hurricanes

A B C D E

24. Potatoes, mice, animals

A B C D E

25. Table ware, objects made of silver, knives

A B C D E

26. Dogs, poodles, animals

A B C D E

27. Trees, tomatoes, grains

A B C D E

28. Shades of red, colors, sizes

A B C D E

29. Pets, rabbits, white animals

A B C D E

30. Baskets, hats, objects made of straw

A B C D E

DO NOT GO BACK TO PART 1 AND DO NOT GO ON TO ANY OTHER TEST UNTIL ASKED TO DO SO

STOP

Appendix G
Percent Correct on the Diagramming Relationships Test

Item	*M*	S.D.
16	.57	.50
18	.34	.47
20	.91	.29
22	.56	.50
24	.89	.31
26	.76	.43
28	.76	.42
30	.47	.50
17	.93	.26
19	.62	.48
21	.57	.50
23	.59	.49
25	.41	.49
27	.66	.47
29	.50	.50

Note. *n* = 401.

Appendix H

Name ____

LETTER SETS TEST -- I-1 (Rev.)

Each problem in this test has five sets of letters with four letters in each set. Four of the sets of letters are alike in some way. You are to find the rule that makes these four sets alike. The fifth letter set is different from them and will not fit this rule. Draw an X through the set of letters that is different.

NOTE: The rules will <u>not</u> be based on the sounds of sets of letters, the shapes of letters, or whether letter combinations form words or parts of words.

<u>Examples</u>:

A. NOPQ ~~DEFL~~ ABCD HIJK UVWX

B. NLIK PLIK QLIK ~~THIK~~ VLIK

In Example A, four of the sets have letters in alphabetical order. An X has therefore been drawn through DEFL. In Example B, four of the sets contain the letter L. Therefore, an X has been drawn through THIK.

Your score on this test will be the number of problems marked correctly minus a fraction of the number marked incorrectly. Therefore, it will <u>not</u> be to your advantage to guess unless you are able to eliminate one or more of the letter sets.

You will be allowed <u>7 minutes</u> for each of the two parts of this test. Each part has 1 page. When you have finished Part 1, STOP. Please do not go on to Part 2 until you are asked to do so.

DO NOT TURN THIS PAGE UNTIL ASKED TO DO SO

Part 2 (7 minutes)

16.	ABCX	EFGX	IJKX	OPQX	UVWZ
17.	LNLV	DTFL	CLNL	HRLL	LLWS
18.	ABCE	EFGI	IJKM	OPQT	UVWY
19.	GFFG	DCCD	STTS	RQQR	MLLM
20.	DCDD	HGHH	MMLM	QQQR	WWVW
21.	FEDC	MKJI	DCBA	HGFE	JIHG
22.	BDBB	BFDB	BHBB	BBJB	BBLB
23.	BDCE	FHGI	JLKM	PRQS	TVWU
24.	BDEF	FHIJ	HJKL	NPQR	SVWX
25.	NABQ	PEFS	RIJV	GOPK	CUWG
26.	FEGF	KLHJ	NOQP	PQSR	TURS
27.	AOUI	CTZR	JHTN	PBRL	RTVH
28.	BEPW	HJTX	KNRZ	KOSV	WRPM
29.	RRBR	QQAR	FTEF	JXIJ	SSCS
30.	QIFB	CGIJ	BCOR	ZRED	JIFC

DO NOT GO BACK TO PART 1 AND

DO NOT GO ON TO ANY OTHER TEST UNTIL ASKED TO DO SO.

<u>STOP</u>.

Appendix I
Percent Correct on the Letter Sets Test

Item	*M*	*S.D.*
16	.92	.28
18	.83	.38
20	.85	.36
22	.90	.30
24	.77	.42
26	.54	.50
28	.17	.38
30	.08	.28
17	.86	.35
19	.80	.40
21	.83	.37
23	.76	.42
25	.62	.49
27	.43	.50
29	.31	.46

Note. *n* = 401.

Appendix J

DEDUCTIVE REASONING TASK

Elliot, Ron, Lois, Nancy, and Fran, whose last names are Adams, Davis, Gordon, Harris, and Jackson, have favorite types of books. The books are mystery, biography, science fiction, comedy, and poetry.

Match everything up from the clues below.

1. Jackson and the person whose favorite is mystery bought Harris his favorite book.
2. Adams spends one hour each night reading his favorite, which isn't biography.
3. Lois, Gordon, and the girl whose favorite is mystery went to the book store with Elliot and Adams.
4. The girl whose favorite is science fiction talked to Lois about the latest best sellers.
5. Adams and the person whose favorite is comedy walked to school with Lois.
6. Fran's favorite is not mystery, but she likes it alot, anyhow.

MATRIX FOR DEDUCTIVE REASONING TASK

Helpful hint: To fill in the chart, make a note in each box if the combination can be eliminated. When there is only one blank box left in the row or column within the category, mark that box as a reminder that you have found one of the solutions. Continue through the clues, eliminating as many of the boxes as you can.

	A	D	G	H	J	M	B	SF	C	P
E										
R										
L										
N										
F										
M										
B										
SF										
C										
P										

Appendix K

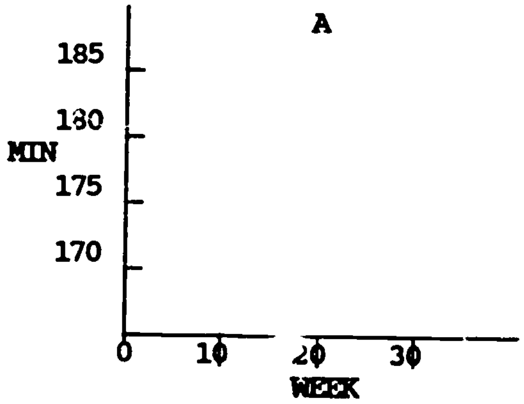
Systems Thinking Instrument
May, 1988

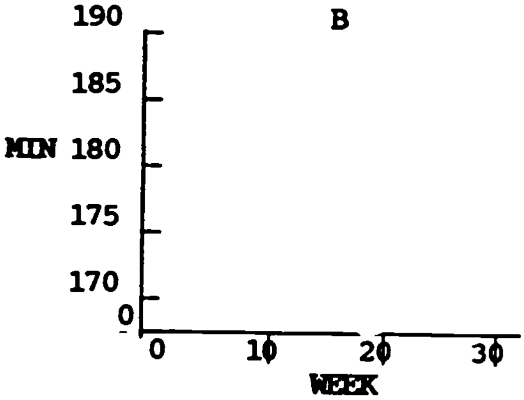
Name____________________ _____ Class ____________________

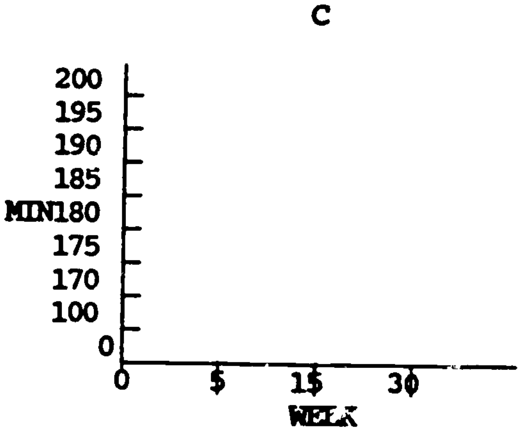
John was training to run the Boston Marathon. To check his progress, he graphed his times each week with the following results:

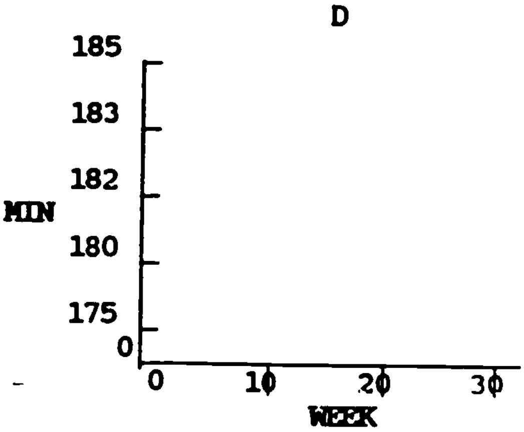
Week	Minutes
1	185
5	180
10	175
15	188
20	185
25	180
30	170

1. Circle the set of axes that best describes the results?





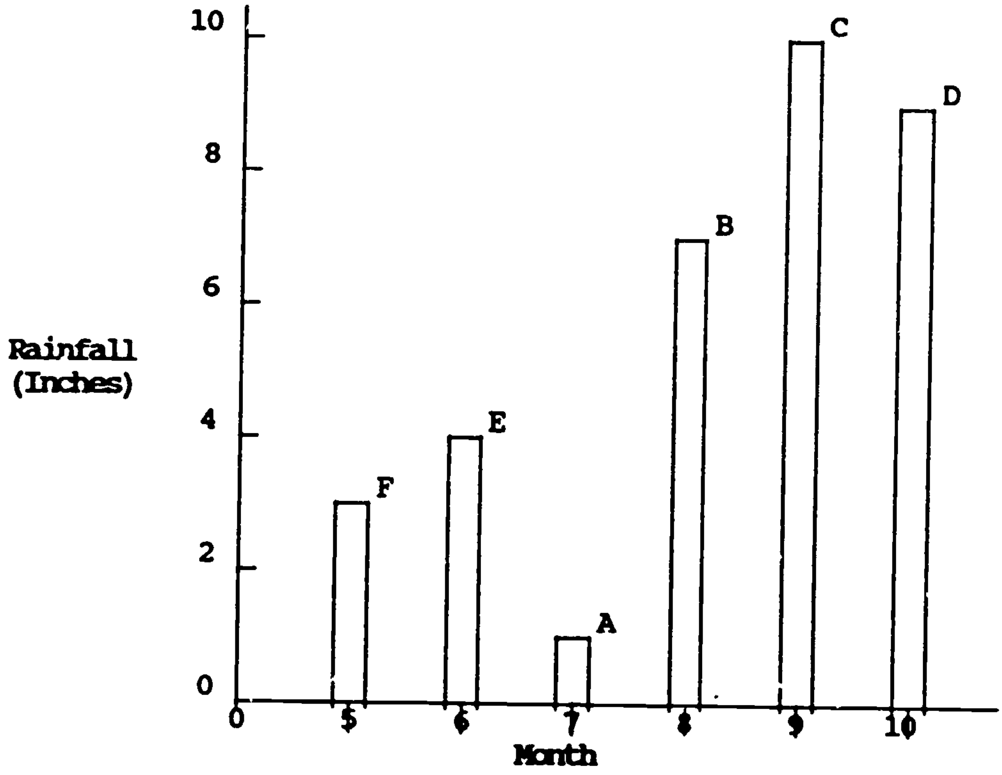




2. Now select the appropriate axes and graph John's training times.

3. Is his training working? Why?

Rainfall was recorded for six months last year. Based on the measurements, here are the number of inches that fell per month.

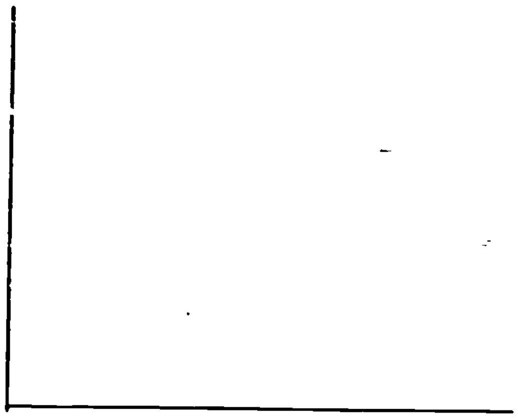


4. What are the coordinates for point C? ______

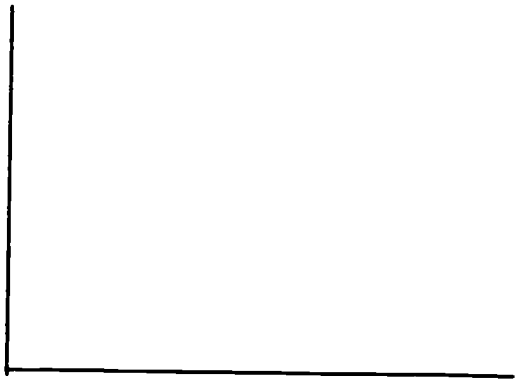
5. During what month did it rain 4 inches? ______

6. How many inches of rain fell during month 8? ______

7. A student prepares for an examination. The more she prepares, the better her performance, up to a point, after which additional study does not improve her test score. Translate the this relationship into a graph.



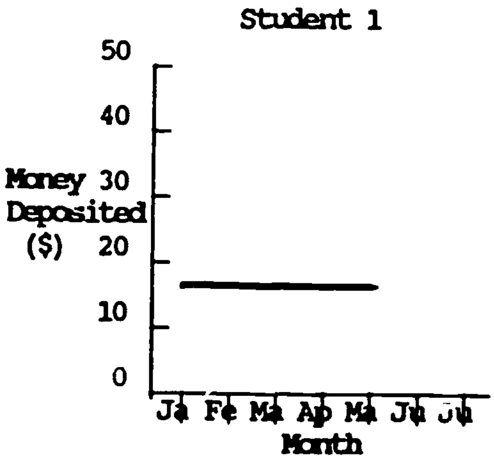
8. Amount of exercise causes changes in weight. Graph the following relationship. As exercise increases, weight decreases. Make sure to label the appropriate axes.

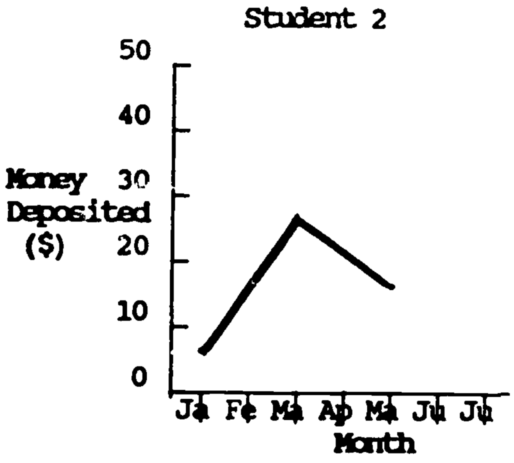


When Susan took physical education, her ability to run the mile steadily improved until she was able to run the distance in 8 minutes. She continued at that rate until the end of the school year. When Susan returned from summer vacation, she found that her speed had dropped to 11 minutes per mile. Soon after and with some practice, her speed quickly returned to 8 minute miles.

9. Graph the pattern of Susan's time for the mile.

10. Two students had part-time jobs after school. Each student deposited a portion of their earnings into a savings account. The following graphs show the pattern of savings for each student from January through May.


Student 1


Student 2

a. How would you describe student 2's pattern of savings?

b. How does student 1's pattern of savings compare with student 2's?

c. Who saved more money in the from January to May? ____________

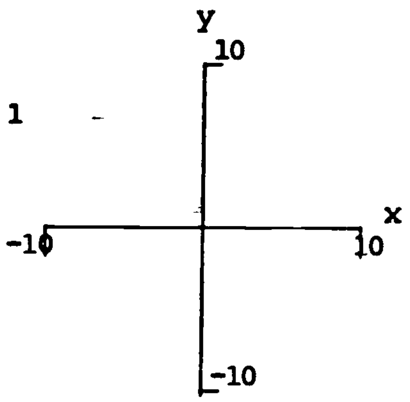
d. How much more money did that student save? $______________

e. Assume each student continued to save at the same rate, plot data points for June and July on the graphs above.

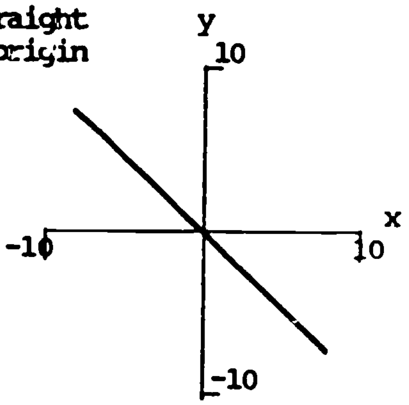
11.

a. Graph the function: $y = 3x - 1$

b. What is the slope of this line?

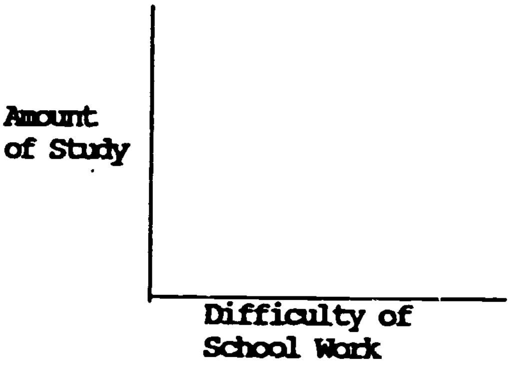


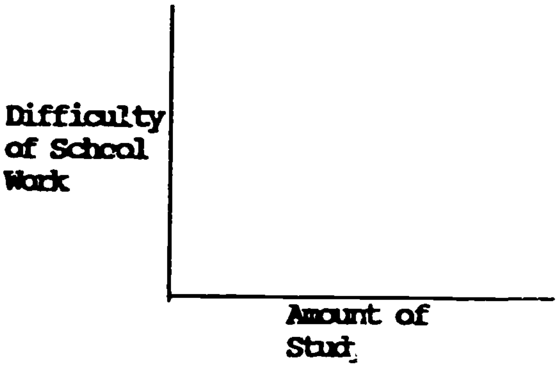
12. Write the function for the straight line that passes through the origin and has a slope of -1.



13.

a. Select one of the sets of axes and graph the relationship between difficulty of school work and amount of study.

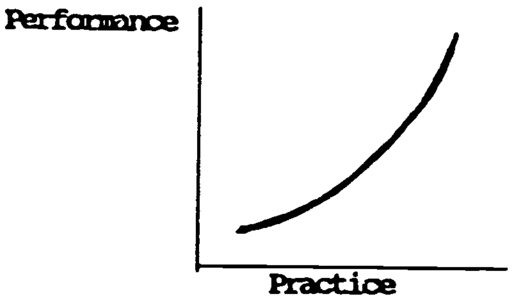


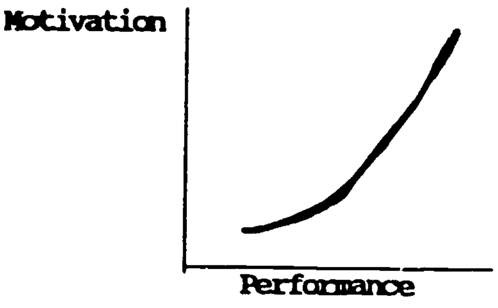


b. Briefly describe the relationship between study time and difficulty of school work (how difficulty affects study time and how study time affects level of difficulty).

c. Show how this relationship can be represented in a feedback loop. Add the + or - to indicate the same (+) or opposite (-) direction of relationship.

14. Describe the relationship between practice and performance, performance and motivation, and among practice, performance, and motivation.





A. Complete the following sentences to describe a causal relationship.
B. Then draw an arrow diagram to represent causation.
C. Next, indicate if the change in the second variable is the same (+) or opposite (-) from the first variable.

EXAMPLE: Consumption of desserts causes <u>changes in weight</u>.

consumption of desserts $\xrightarrow{+}$ weight

15. Amount of exercise causes ____________________.

exercise

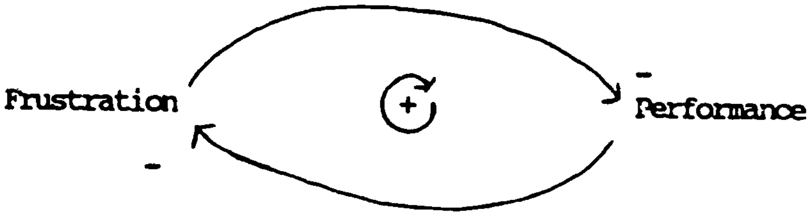
16. Number of births causes ____________________.

number of births

17. The amount of study time causes ________________.

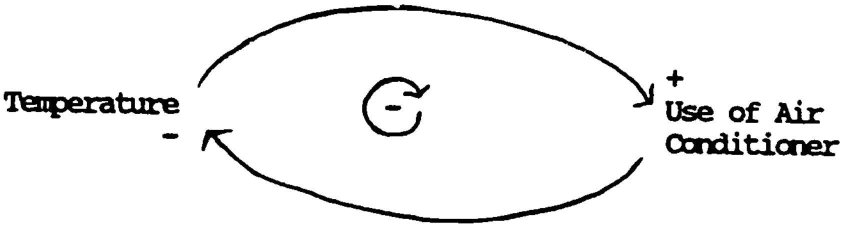
study time

18.



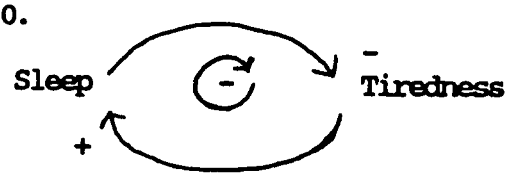
Describe the relationship between frustration and performance.

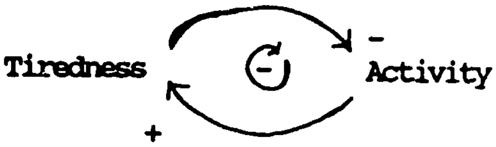
19.



Describe the relationship between temperature and use of an air conditioner.

20.





Describe these two feedback loops.

21. For each set of three variables, state whether it is a stock (level) or a flow (inflow or outflow) (rate). Give the unit of measure for each variable.

Example:

Variables	Stock or Flow	Unit of Measure
Population	Stock	Number of people
Births	Flow (Inflow)	Babies born per year
Deaths	Flow (Outflow)	Deaths per year

a. Credits
   Bank Balance
   Debits

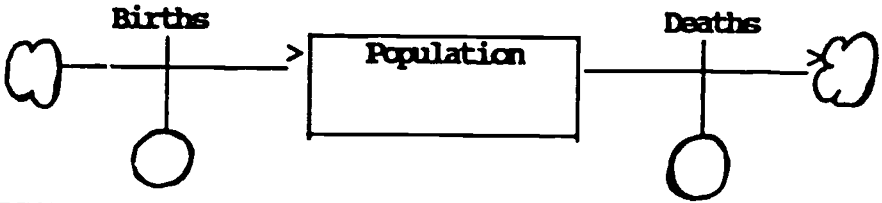
b. Evaporation
   Rainfall
   Water

c. Completions
   Assignments
   Schoolwork

d. Inventory
   Sales
   Production

22. For the following set of variables for flight patterns, create a model using a flow diagram. Briefly explain how the model works.

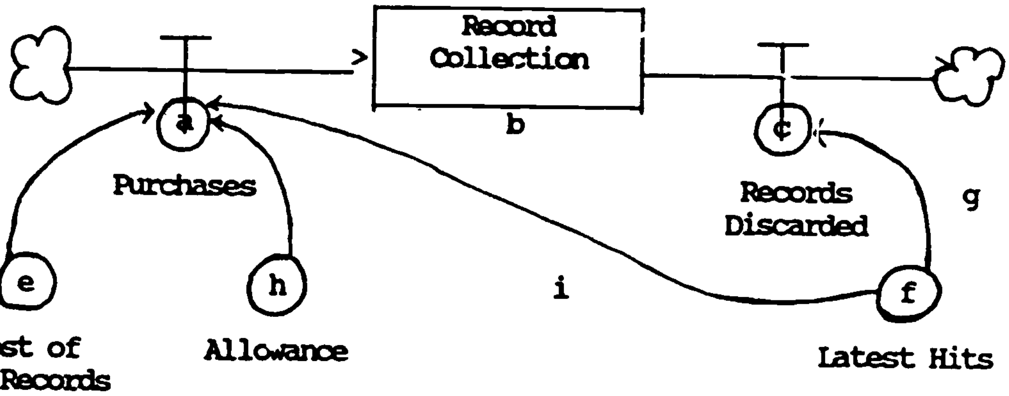
EXAMPLE: Population, Deaths, Births



FLIGHT PATTERN VARIABLES
Departures
Planes
Arrivals

23. Using the flight pattern model above, what variable(s) might affect the number of departures?

24. Here is a model of a student's record collection. Label the numbered parts of this model as a flow, stock, converter, or connector.



Label the following:

a. b. c.
d. e. f.
g. h. i.

25. <u>According to the model</u>, what influences the amount of records in the collection?

26. <u>According to the model</u>, what influences the number of purchases?

27. How might the purchase of a compact disc player affect the model?

28. Each year the bluewinged finch must make the journey from South America to Iceland in time for mating season. The birds leave South America and stop only once (in Maryland) to restock their food supply. They must rely on sand crabs as their only food source. After gorging themselves, those birds that are able to fly (usually a small percentage are unable to make the trip) continue on to Iceland, where they then mate and eventually return to South America. Each finch couple produces two babies per year.

a. What variables might affect the birth rate of the bluewinged finch?

b. Draw a flow diagram to show the mating system of the bluewinged finch.

c. During the fifth year of migration, a disease infected the population of crabs, causing a severe food shortage. The number of crabs slowly returned to normal by the tenth year. This ten year cycle recurred. Draw a graph to show what would happen to the finch population as their food source increases and decreases.

Appendix L

ITEM STATISTICS FOR THE SYSTEMS THINKING INSTRUMENT

Item	Maximum Score	GPS		Biology		Chemistry		Physics		War & Rev.	
		M	S.D.	M	S.D.	M	S.D.	M	S.D.	M	S.D.
1	1	.74	.45	.88	.33	.68	.50	.72	.45	.78	.44
2	2	1.56	.78	1.55	.64	1.62	.69	1.80	.49	1.78	.44
3	2	1.24	.74	1.48	.79	1.50	.81	1.91	.35	2.00	.00
4	1	.41	.50	.57	50	70	.46	.89	.32	.78	.44
5	1	.91	.29	.94	.24	.97	.18	.98	.14	1.00	.00
6	1	.94	.24	.97	.17	.97	.18	1.00	.00	1.00	.00
7	3	1.59	1.28	[illegible]	1.04	2.38	1.01	2.59	[illegible]	2.33	1.00
8	3	1.65	1.01	2.11	1.03	2.23	.96	2.80	.56	2.56	.88
9	3	.97	.94	1.02	.86	1.32	.79	1.39	.98	1.89	1.05
10A	3	1.50	1.16	1.55	.95	1.87	1.13	2.00	.95	1.78	.97
10B	3	1.88	1.22	2.22	1.12	2.15	1.15	2.17	1.16	2.00	1.12
10C	2	1.35	.95	1.38	.93	1.57	.83	1.56	.84	1.56	.88
10D	1	.15	.36	.42	.50	.42	.50	.54	.50	.67	.50
10E	2	1.12	.88	1.43	.75	1.45	.72	1.56	.60	1.67	.50
11A	2	.09	.29	.57	.86	.80	.94	1.30	.84	1.44	.88
11B	1	.03	.17	.28	.45	.42	.50	.76	.43	.89	.33
12	2	.06	.34	.48	85	.87	.98	1.15	1.00	1.78	1.67
13A	2	.32	.68	.66	.85	1.32	.85	1.39	.90	1.00	1.00
13B	4	1.32	1.30	2.48	1.23	2.75	1.24	2.50	1.45	2.22	1.64
13C	5	1.20	1.34	3.85	1.55	2.30	1.87	2.85	1.63	4.00	1.32
14	4	1.47	1.64	2.46	1.76	2.17	1.70	2.63	1.62	2.33	1.73
15	3	2.20	1.15	2.74	.62	2.63	.60	2.74	.65	3.00	.00
16	3	2.03	1.29	2.85	.43	2.55	.87	2.72	.68	3.00	.00
17	3	2.26	1.16	2.66	.73	2.57	.83	2.74	.65	3.00	.00
18	4	1.76	1.33	2.83	1.44	2.60	1.48	2.94	1.58	3.11	1.76
19	4	1.76	1.02	3.06	1.18	2.73	1.35	2.94	1.43	2.67	1.66
20	4	1.91	1.24	2.71	1.23	2.63	1.28	2.63	1.52	2.89	1.69
21A1	1	.32	.47	.71	.46	.67	.48	.70	.46	.78	.44
2	1	.06	.24	.34	.48	.23	.43	.28	.45	.33	.50
3	1	.35	.48	.74	.44	.82	.39	.89	.32	1.00	.00
4	1	.15	.36	.58	.50	.60	.49	.67	.48	.78	.44
5	1	.33	.48	.71	.46	.70	.46	.72	.45	.78	.44
6	1	.09	.29	.35	.48	.28	45	.28	.45	.33	.50
B1	1	38	.49	.83	.38	.82	.39	.87	.34	1.00	.00
2	1	.18	.39	.40	.49	.25	.44	.35	.48	.44	.53
3	1	.38	.49	.74	.44	.75	.44	.83	.38	1.00	.00
4	1	.20	.41	.51	.50		.44	.33	.48	.44	.53
5	1	.38	.49	.77	.42	[illegible]	.40	.87	.34	1.00	00
6	1	.15	.36	.51	.50	.48	.50	.61	.45	.57	.53
C1	1	.24	.43	.83	.38	.63	.49	.85	.36	1.00	.00

- CONTINUED -

Appendix L Continued

ITEM STATISTICS FOR THE SYSTEMS THINKING INSTRUMENT

Item	Maximum Score	GPS		Biology		Chemistry		Physics		War & Rev.	
		M	S.D.	M	S.D.	M	S.D.	M	S.D.	M	S.D.
2	1	.06	.24	.46	.50	.52	.50	.56	.50	.67	.50
3	1	.24	.43	.77	.42	.57	.50	.85	.36	1.00	.00
4	1	.09	.29	.43	.50	.48	.50	.46	.50	.44	.53
5	1	.32	.47	.77	.42	.70	.46	.87	.34	1.00	.00
6	1	.03	.17	.31	.46	.37	.49	.52	.50	.78	.44
D1	1	.24	.43	.58	.50	.60	.49	.83	.38	.89	.33
2	1	.03	.17	.35	.48	.28	.45	.61	.49	.67	.50
3	1	.24	.43	.68	.47	.67	.48	.83	.38	.89	.33
4	1	.06	.24	.37	.49	.33	.48	.57	.50	.56	.53
5	1	.24	.43	.58	.50	.58	.50	.82	.39	.89	.33
6	1	.03	.17	.34	.48	.28	.45	.5[illegible]	.50	.56	.53
22	7	2.65	2.06	4.12	1.62	4.12	1.92	4.35	2.19	5.78	1.39
23	2	.76	.78	1.43	.73	1.42	.81	1.56	.69	1.78	.44
24A	1	.74	.45	.68	.47	.73	.44	.91	.29	1.00	.00
24B	1	.76	.43	.91	.29	.88	.32	.91	.29	1.00	.00
24C	1	.70	.46	.69	.46	.72	.45	.89	.32	1.00	.00
24D	1	.35	.48	.85	.36	.85	.36	.91	.29	1.00	.00
24E	1	.32	.47	.78	.41	.87	.34	.89	.32	1.00	.00
24F	1	.32	.47	.82	.39	.80	.40	.89	.32	1.00	.00
24G	1	.32	.47	.85	.36	.85	.36	.91	.29	1.00	.00
24H	1	.41	.50	.7[illegible]	.43	.87	.34	.89	.32	1.00	.00
24I	1	.38	.49	.85	.36	.82	.39	.91	.29	1.00	.00
25	3	1.20	.84	1.71	.84	1.85	.88	1.72	.90	2.11	.78
26	2	1.35	.88	1.88	.41	1.82	.57	1.82	.58	2.00	.00
27	3	.82	.97	1.46	1.17	1.77	1.14	1.50	1.00	2.22	.97
28A	5	1.00	.82	1.91	1.10	2.02	1.11	2.11	1.25	2.33	1.66
28B	12	1.24	1.56	2.22	2.39	1.28	1.64	2.52	2.44	2.44	2.35
28C	[illegible]	.30	1.05	.95	1.38	.72	1.33	1.46	1.68	2.11	1.54
Graph	6	4.56	1.31	4.91	1.06	4.93	1.19	5.39	.79	5.33	.50
G.Interp	19	8.91	3.80	11.98	4.30	12.42	3.51	13.30	3.71	12.56	3.88
G.Trans	17	6.15	3.87	8.29	3.20	9.42	3.03	11.19	3.03	11.56	3.05
G.Math	5	.18	.58	1.32	1.99	2.08	1.97	3.20	1.84	4.11	1.69
Loops	17	6.65	3.74	12.45	4.12	10.27	4.48	11.37	5.17	12.67	6.18
Causality	9	6.50	3.36	8.25	1 63	7.75	2.40	8.20	1.58	9.00	.00
Ident	33	9.12	8.59	20.83	8.89	20.07	8.18	23.85	7.88	26.78	4.68
Sys.Th	34	9.03	5.30	14.72	5.50	14.27	4.81	15.57	5.89	18.67	2.60
Total	140	51.09	22.63	82.75	21.67	81.20	19.00	92.07	20.11	100.67	1[illegible].83

Note: $n_{GPS} = 34$ $n_{BIO} = 65$ $n_{CHEM} = 60$ $n_{PHYS} = 34$ $n_{W \& R} = 34$

Scores for the two War and Revolution students concurrently enrolled in chemistry are included in the chemistry and War and Revolution means

Appendix M

Prepared by M. E. Thorpe

Curriculum and Instruction

The science teachers integrated the systems thinking approach into instruction in distinct ways. Because of variation in how the systems thinking approach was integrated into instruction, it is useful to examine each subject separately. For all content areas, the objectives and general approaches to instruction are presented, followed by a discussion of the content and methods used in teaching with a systems perspective. Instructional activities, requisite cognitive skills, and content knowledge are described. Comparisons between traditional and systems instruction focus on information related to topics taught with the systems approach.

General Physical Science (GPS)

Systems instruction was used to teach speed, motion, velocity, and acceleration, concepts traditionally difficult to master. Although students learn formulas and terminology, few have an intuitive understanding of relationships underlying these ideas. To promote comprehension, the teacher used an inquiry-oriented format and a combination of traditional and systems methods that was supplemented with structural models.

Sequence and organization of systems instruction. Systems initially was taught through exploration of models and their use to solve problems. Drawing on examples from the video "Search for

Solutions" and actual experiences, students described problems and identified elements of models. To reinforce the notion of causality, students were exposed briefly to causal loop diagrams and asked to construct loops of scientific and non-scientific problems. Structural diagramming was introduced, in conjunction with the Macintosh, as another modeling tool.

Instruction began with a demonstration of a water tank model using STELLA. Students worked online in pairs to build and explore structural models using worksheets prepared by Technical Education Research Centers (TERC). Students experienced difficulty with the task procedurally and declaratively. Although directions were given to guide students through a sequence of steps, the objective of the activity was not stated, leaving many students uncertain about what they were doing and why. Another problem was the lack of familiarity with the Macintoshs and STELLA, compounded by operational mechanics and the need to alternate among different representations.

The teacher subsequently adopted a more direct approach. He presented a simpler offline model for discussion with the whole class. Students were shown how the concept of rates was represented in a structural model. Students then worked through problems together, creating graphs of expected rates and flows.

After a month's interim, systems thinking was reintroduced to teach speed, distance, acceleration, and velocity, with augmentation by traditional methods. Students spent a week on offline problems using formulas. They also were required to describe the motion of objects in verbal or graphical form.

Students then reviewed STELLA symbols, discussed a model of acceleration, velocity and distance and worked through problems of stocks and flows. They worked online for one week creating structural models. Following detailed instructions they built simple diagrams of velocity (a rate) and distance (a stock). They ran models, changed parameters, and explained results. They also used their models to solve traditional word problems.

A more complex model was introduced to illustrate relations among acceleration, velocity, and distance. Students were given a structural diagram and parameters for each variable and were directed to create graphs, change values, predict, and describe the observed behavior. They were able to do only a few problems and many had difficulty generating and interpreting the graphs. An acceleration lab then was conducted. Students watched a demonstration, collected data, and graphed their results. The class then compared the STELLA models with other representations of speed and motion.

Assessment of systems instruction. Speed and motion usually are difficult for students to learn due to the concept of rates. The systems approach presents rates structurally and graphically, allowing students to define rates and relationships in a model. Additionally, students can create structural diagrams and illustrate how the underlying behavior changes over time. When students are instructed with traditional methods. they often fail to grasp that rates change over time. They focus on mastering formulas. Although they might compute the correct answer, they often cannot explain or predict the behavior over time.

While systems instruction helped some students, others had difficulty understanding speed and motion. Understanding acceleration as a rate of change of a rate of change was difficult even with systems. Some students had trouble with the logic underlying velocity's representation as both a stock and a flow. Furthermore, many students initially had difficulty with the mechanics of STELLA until the teacher provided explicit directions and sample models. The teacher expressed disappointment with students' explanations of observed behavior, despite atypical task demands. Increased exposure to this type of task, as well as more examples of process analysis may facilitate performance.

Comparison with traditional instruction. The systems and traditional teachers wanted students to attain a basic understanding of speed and motion. Although they shared this goal, they differed in their approaches to instruction and the amount of time devoted to the topic. The traditional teacher devoted six weeks to the unit; the systems teacher spent four weeks. The traditional teacher emphasized definitions, measurement, and formulas. She closely followed the text, emphasizing a computational approach to speed and motion problems. Students were instructed to write formulas, enter the numbers into the formulas, make conversions, and check their answers for accuracy. In contrast, the systems teacher stressed concepts underlying the formulas. Although students were exposed to the formula method, instruction was presented primarily from a systems perspective.

Biology

In biology, systems was used to present an integrated view of biological processes. This integrated approach contrasts with traditional instruction which typically emphasizes the functions and characteristics of biological processes in a more segmented way. In comparison to the other sciences, systems instruction in biology was used primarily to reinforce and/or synthesize information presented though other approaches.

Systems was applied differently in the instruction of cellular transport and cellular response to infection. With transport, systems was used to reinforce concepts students had learned through other activities. They worked on several labs related to cellular transport prior to the systems module. The systems model was a simulation of their lab experiences. Systems was used to synthesize and extend information previously taught in immunology. Although students spent several weeks studying characteristics, functions, and selected relationships among the components of the immune system, they had not discussed the way these components operate as a system. The model was the first time students were exposed to how these components interact during cellular response to infection.

Sequence and organization of systems instruction. The systems teacher used a direct approach to instruction. In comparison to the other science classes, more time was devoted to whole class presentations and discussion. He spent approximately one week introducing students to modeling and STELLA. As in the other sciences, students were informed that models could be used as a problem solving tool in many subject areas. Although causal models were introduced, greater emphasis was placed on teaching systems from a structural perspective.

Students were introduced to structural diagrams and STELLA through demonstration and discussion of a population model. They examined the components of the model and how factors affected behavior over time. They also made predictions about patterns of behavior and altered them based on changes in the model. Graphs were examined with respect to predictions. To demonstrate STELLA's simulation capacity, the next model was done in conjunction with a lab on the cooling rate of liquid. At fixed intervals, the teacher reported the temperatures of three soup containers of equal volume that differed in initial temperature. Students recorded data and created graphs of temperature changes over time. The teacher then demonstrated a structural model of the experiment. Students compared their graphs to the one generated by STELLA. As a class they suggested modifications to the parameters of the model and observed changes in cooling rate.

Students were instructed to operate the computers through the software program, Guided Tour of the Macintosh. After completing this program students, worked online in pairs to explore a teacher-constructed model of oxygen production during photosynthesis. The model illustrated the relationship between light intensity and the rate of oxygen production. Although the teacher had introduced the model previously in class discussion, students had not studied this topic before, and therefore, they had difficulty understanding relationships. The use of STELLA facilitated students' subsequent performance in modeling and understanding of the concept.

Following a series of learning activities on the topic of diffusion and osmosis, a structural model was introduced to reinforce these processes. In this instance, the teacher led the class to construct a complex structural model containing two interrelated subsystems, one for osmosis and another for diffusion. The students and teacher indicated that the process of group modeling was helpful to understand the relationships in the model. Students then manipulated this model online. Exercises guided students to examine and predict the behavior of factors over time as well as the relationships among factors in the model. The module was summarized by reviewing the STELLA graphs.

After a four month interval, students were reintroduced to systems instruction through causal loop diagramming, beginning with single loop structures and progressing to increasingly complex multiple-loop diagrams. These lessons were d[illegible] in preparation for causal modeling during discussion of the immune system. The teacher devoted a month to the immune system. He reviewed characteristics and functions of cells and proteins, how they combat infection, and different types of immunity through traditional instructional approaches. Instruction emphasized the causes of disease, how disease is spread, and the body defends against infection. Students learned techniques to isolate and study the characteristics of bacteria from lab activities. They also examined the effects of different substances on controlling the spread of microbes.

Following this substantive preparation, the teacher guided the class to construct a causal loop diagram of cellular response to infection as a means of integrating and extending knowledge of the immune system. The model was constructed in successive steps as students were prompted to suggest and defend connections among factors. Students made predictions about the behavior of the immune response and the effects of new connections on various parts of the system. The model was the first time students were exposed to a systemic view of cellular response to infection.

Directly related to the immune system was the model of AIDS which the MIT consultant presented as a causal loop. The consultant presented the model and conducted a class discussion on the spread of AIDS, the model's structure and components, and policy decisions related to the disease.

Assessment of systems instruction. Observations and anecdotal reports indicated a high level of student engagement during systems instruction. Group modeling as an instructional vehicle challenged and engaged students to participate in these discussions. During class discussions students were able to identify important factors within biological systems, suggest and explain cause-and-effect relationships, and predict outcomes. However, results from transport and immunology tests were disappointing. Students indicated that they were confused by certain questions and that they needed additional instruction in certain areas particularly graphing.

Comparison with traditional instruction. Although the systems and traditional teachers used the same text, they differed in the amount of time devoted to topics. The systems teacher spent five weeks on both transport and immunology; the traditional teachers devoted three weeks of instruction to each topic. There also were differences in the number and type of activities used to supplement the text. Differences in pedagogical approach were observed although the three teachers used class discussion as the primary method of instruction. One traditional teacher proceeded slowly and had difficulty with classroom management. The other traditional teacher used various strategies to promote student understanding and retention of information. Through questioning, visual cues, and analogies, she directed students to connect prior knowledge and new information. Her presentations and discussions were clear, well organized, and fast-paced. The systems teacher shared some characteristics of both traditional teachers. His lessons were organized and he used visual materials effectively. He tended to elicit students' ideas rather than to supply responses. His class was attentive although the pace of his lessons was slower and more similar to the first traditional teacher.

Chemistry

The systems approach was used to augment instruction on reactions and provide an analytic tool to promote thinking about cause-and-effect relationships and behavior change. Although reaction rates is not a difficult topic, the teacher anticipated that systems might broaden understanding and provide a problem solving tool that could be applied to this and other chemical concepts. The selection of reaction rates as a systems module represented an ideological shift on the part of the teacher. His initial goal, to examine chemically related social problems (e.g. pollution), proved unsatisfactory because the chemistry underlying the models was too complex. Furthermore, the social problems were removed from the core curriculum. The teacher therefore sought an alternate approach to integrate systems instruction into existing course content. However, identifying appropriate instructional topics was difficult. Instead he modified instructional activities developed last year and focused systems instruction on rates of chemical reactions.

Sequence and organization of systems instruction. Students were introduced to systems through causal loop diagrams because models could be constructed from verbal descriptions without the complexity of quantification. Students described patterns of behavior over time and examined connections among variables, cause-and-effect relationships, interactions, and feedback.

Three models were introduced: the rate at which dough rises; reactants and products of the nitrogen cycle; and burnout. The goal was to model these processes and describe the relationships among factors. Demonstrations, guided discussion, and assignments were used to help students construct causal models.

The topic of reaction rates was introduced with traditional methods. Students learned to apply formulas to express and compare reaction rates. Although systems models had not yet been constructed, students were exposed to concepts related to system instruction. They conducted experiments and discussed the effects of alterations in factors such as temperature and concentration on reaction rates. They also constructed graphs of lab results to view changes in reactants and products over time.

During the second week of instruction the teacher gave a demonstration of modeling a reaction with STELLA. Students then explored STELLA following a prepared script to construct a chemical reaction model with predefined connections. This lesson provided guided exploration of the operation of the Macintosh and STELLA's structural and graphical features. Students worked online for two weeks developing and altering reaction models. Models were based on previous labs, enabling students to apply prior knowledge to develop hypotheses of expected outcomes. Online exercises required students to construct, alter, and test models with different parameters, then predict, describe, and explain the behavior. A final exercise required students to construct models to see if they could apply generic models to novel problems outside of chemistry.

Assessment of systems instruction. The teacher expressed mixed reactions to systems instruction. He felt he spent too much time on reaction rates at the expense of other topics, especially given that the concept is not critical for high school chemistry. The topic of rates was selected because it was amenable to the approach, and therefore offered an opportunity to illustrate how systems could be applied to instruction in chemistry. However, instruction exceeded conventional coverage of this topic due to the time was needed to provide both traditional and systems activities, particularly student modeling.

Learning STELLA was difficult for some students. They had problems understanding how to specify and operate their models both mechanically and logically. The reaction rate formula was easier to apply the operation of STELLA. On the positive side, the teacher indicated that construction of reaction rate models facilitated students' understanding of chemical equilibrium, a traditionally complex concept. Students understood equilibrium more readily than they had in the past. Furthermore, understanding equilibrium assisted students to learn topics such as acids and bases and solubility of products.

STELLA was generally perceived as a useful supplement to lab activities. The teacher indicated that although labs were helpful to develop conceptual understanding, students sometimes were confused by the procedures and their efforts to find the right answer. Moreover, labs did not always work due to imprecise methods. Thus, the potential benefits of labs were reduced when students did not see the desired outcomes. STELLA could replicate observed lab experiences and simulate other conditions without need to repeat the experiment.

Comparison with traditional instruction. Both the systems and traditional students covered many of the same chemistry concepts with different instructional activities, emphases, and exposure. Traditional classes emphasized mastery of algorithms in contrast to developing models to define relationships between concepts. Under typical conditions, the systems and traditional teachers had similar instructional styles. Instruction was moderately paced and marked by frequent questioning. Both teachers used a lecture-discussion format and communicated in clear, concise terms. They also used the similar lab activities. For the study of reaction rates, the systems teacher departed from his usual approach. Instead of a using direct instruction, he organized student activities to promote inquiry. Students worked online for two weeks. The teacher served as a facilitator rather than an instructor, responding to questions and providing help as needed.

Physics

Systems thinking was used to augment several physics units, including motion, momentum, energy, and electricity. The modules exposed students to both traditional formulas and the systems approach. STELLA was used to develop a deeper understanding of relationships among physical concepts and observe changes in behavior over time.

Sequence and organization of systems instruction. Parallels between instruction in physics and GPS were found, particularly the emphasis on structural modeling. Although causal loops were introduced briefly, structural diagramming was the primary problem solving tool. STELLA was introduced with a model of water tank drainage. The teacher demonstrated how STELLA operated while discussing the physical relationships displayed in the model. Students then spent two weeks building structural models of motion. The introduction of STELLA was facilitated by a number of students who remembered its mechanics from systems chemistry.

Textbook exercises that required mathematical formulas, provided opportunities to solve problems and compare traditional and systems solutions. Typically students worked on these problems for several days then discussed the results. STELLA was used to model problems illustrating the conservation of momentum. Students were introduced to a general model of momentum and then given online problems. Using guided inquiry, the teacher helped students to focus on important elements in the problem and appropriate solution processes.

Students spent three days building a complex model of rocket propulsion. The teacher reported that students were able to construct the model with varying degrees of prompting and understood the numerous factors. At the end of the year, the teacher returned to the water tank problem. Students conducted a lab, collected data, drew graphs to illustrate the behavior observed, and then developed structural models. The teacher summarized their experience, reviewed the logic of the model, and discussed other ways they might have solved the problem.

Assessment of systems instruction. The teacher was more satisfied with systems instruction in physics than GPS. Physics students were better able to cope with the applications and mechanics of modeling. Nevertheless, some of the same difficulties found in GPS were noted in physics. Students differed in their need for help and tolerance for ambiguity with instruction through guided inquiry. Some students had problems with the operation of STELLA. Common problems involved defining graphs and initializing values. Some physics students had trouble conceptualizing problems as diagrams, confusing stocks and flows. Others did not approach problems analytically, failing to use knowledge of physics to predict what their structural diagrams should look like. As a consequence, some students had difficulty interpreting graphs and altering their model based on their understanding of the phenomena.

The teacher noted that when questioned individually to focus on key elements of the problem, students responded appropriately. More exposure to generic models might be an effective instructional tool. Moreover, some students might benefit from explicit instruction on problem solving strategies including self-regulation techniques to direct attention to important problem elements. In contrast, a large proportion of students understood how to use STELLA. Encouraged by the experience, the physics teacher will expand his use of the systems thinking approach next year.

Summary

Objectives of instruction. Systems thinking was used to facilitate understanding of the behavior of scientific phenomena over time. Teachers expressed dissatisfaction with current instructional methods, viewing traditional approaches as developing fragmented knowledge. The approach was used to promote integrated knowledge and develop analytic skills. Although the immediate goal was to improve knowledge acquisition, the teachers expected that skills developed by the systems approach might generalize to other disciplines. Thus, systems was used to augment instruction and teach heretofore difficult topics and concepts.

Approaches to systems instruction. Although the teachers shared goals and concerns, they differed in their approaches to systems instruction. Content differences and objectives determined how the systems approach was applied in the four courses. For example, systems was used in biology to synthesize and apply prior knowledge to broader problems. In contrast, the goal in physics was to depart from the traditional mathematical approach, reduce the emphasis on mastery of computational formulas, and promote understanding and application of concepts underlying these formulas. The systems thinking approach, used as an analytic problem solving tool, was seen as a way of achieving that goal.

Pedagogical style and the method by which the systems approach was applied determined how instruction was organized. In biology, direct instruction was used to teach causal and structural models. Modeling was taught through teacher-led discussions, supplemented by online exercises to explore and manipulate teacher-constructed models. The other science teachers used inquiry methods to augment structural modeling. Computers were used as an instructional tool on which students constructed and manipulated structural models.

Time constraints for instruction and curriculum development continued to be a concern for the systems teachers. Last year, teachers indicated that too much time was spent on systems modules, at the expense of other topic areas. The chemistry and GPS/physics teachers consequently shortened their systems units. Teachers needed to attain a balance between implementing the systems approach and meeting externally imposed achievement standards. Time also was a factor in curriculum development. The demands of teaching in conjunction with development of new curricular approaches and materials placed heavy demands on teachers' time.

Effects on teaching and learning. Teachers reported that the systems approach heightened their awareness of instructional content and goals. It also provided new opportunities for ongoing student assessment, particularly because STELLA produces representations of the processes and products of learning. STELLA makes it possible to diagnose conceptual deficiencies and strengths from the cognitive representations built in the environment.

The teachers reported that systems is a promising tool because it represents a unique approach to science instruction. Of particular importance is that the approach deemphasizes facts and algorithms, focusing instead on the process of understanding how things work. By constructing and manipulating models, students develop an intuitive understanding of the systemic and dynamic nature of phenomena.

Systems thinking is an approach to teaching and learning activities that is atypical in its emphasis on process rather than only the products of instruction. The focus on solution process is likely to be an unfamiliar and antithetical approach, given pressure for achievement. It is critical that teachers communicate this emphasis to their students and structure instruction to reinforce the pedagogical perspective. Some students, however, will need more guidance to deal with the approach because they are accustomed to selecting rather than generating correct answers.

The systems approach promotes active participation from students and teachers. Teachers demonstrate the processes and logic that underlie model building as an analytic tool. They simulate cognitive approaches to problem solving and demonstrate modeling as an iterative process. Hypothesize-test-reformulate sequences force students to refine representations of systemic phenomena. Thus, the iterative nature of modeling provides feedback that is lacking in much traditional instruction.

Implementation also requires a willingness to experiment with traditional approaches. While all teachers expressed a willingness to experiment and devoted a substantial amount of time to curriculum development, there was variation among the teachers in how systems instruction was implemented. Emphasis was placed on having students think about the nature of problems by viewing interrelationships among factors. They were asked to generate explanations rather than merely identify, define, or compute solutions. Differences were evident among teachers in the use of STELLA, the type and complexity of problems presented, and the manner in which assistance was provided. The variation appeared to be attributable to a knowledge of systems theory, comfort with experimentation, and instructional style. In two cases systems instruction was implemented as a distinct change in the organization and function of the classroom. Students spent considerable time developing and experimenting online with models. These experiences were structured to engage students actively in the tasks while the teacher served as a facilitator, thus, representing a departure from traditional student-teacher roles. In another case, systems was applied more traditionally. The teacher presented systems concepts largely through class instruction in which he sought to shift students' thinking through directed questioning. A fourth teacher was severely impeded in his ability to implement systems instruction due to lack of familiarity with STELLA and the Macintosh.