DOCUMENT RESUME

ED 297 718 IR 013 423

AUTHOR	Mandinach, Ellen B.
TITLE	The Cognitive Effects of Simulation-Modeling Software and Systems Thinking on Learning and Achievement.
PUB DATE	Apr 88
NOTE	19p.; Paper presented at the American Educational Research Association (New Orleans, LA, April 5-9, 1988).
PUB TYPE	Reports - Research/Technical (143) -- Speeches/Conference Papers (150)
EDRS PRICE	MF01/PC01 Plus Postage.
DESCRIPTORS	*Cognitive Processes; Cognitive Tests; *Computer Simulation; Computer Software; Instructional Effectiveness; Instructional Innovation; *Intermode Differences; Models; Pretests Posttests; *Problem Solving; *Science Curriculum; Secondary Education; *Systems Approach; Transfer of Training
IDENTIFIERS	*Systems Thinking Approach

ABSTRACT

This paper examines the effect of using a systems thinking approach in existing secondary school curricula to teach content-specific knowledge as well as general problem solving skills, and the effect of using STELLA (Structural Thinking Experimental Learning Laboratory with Animation), a simulation-modeling software program, as a tool by which to teach systems dynamics and content knowledge. Subjects were secondary students in three general physics, four biology, three chemistry, and one history class taught under the systems approach, and an equal number taught using traditional methods. Pretests and posttests were used to identify subjects' ability, content-specific knowledge, and knowledge of systems thinking. The results indicated that students in the more advanced courses (biology and chemistry) performed better on the systems thinking and computer-based activities. The physical science students performed well on the measurement-related problems, but did less well on the general problem solving or modelling-oriented exercises. Accordingly, these students were the least likely to generalize the systems skills beyond their course. The results raise questions about the applicability of the approach for different subjects and groups of students, and the question of why the impact of the innovation was more apparent in biology and chemistry must be considered. Seven tables supplement the text. (10 references) (EW)

The Cognitive Effects of Simulation-Modeling Software and Systems Thinking on Learning and Achievement

Ellen B. Mandinach
Educational Testing Service

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 using the systems approach in existing secondary school curricula to teach content-specific 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 teach system dynamics and content knowledge. The research focuses on the learning outcomes and transfer that result from using such an approach and software in classroom settings. Comparisons are drawn between traditionally taught courses and those using the systems approach to discern the treatment's impact on learning, teaching, and instructional activities.

Curriculum innovations are never easy to implement nor are they easy to examine systematically. The introduction of microcomputers into classrooms has generated a myriad of such innovations. While teachers struggle to contend with technological advances and their potential implications for instructional and classroom procedures, researchers attempt to document the cognitive, affective, and social effects of the innovations. Both practitioners and researchers are trying to keep pace with and identify ways of functioning effectively within the technological revolution. Teachers may adopt a new piece of software or pedagogical approach if the potential payoff seems worthwhile, whereas researchers may seize the opportunity to study the innovation even if the instructional machinery is not yet in place. A delicate balance must be struck, with teachers and researchers working toward a mutual understanding of the factors, constraints, and perspectives under which the other must function (Mandinach & Thorpe, 1987b). The STACI Project, which examines the impact of one such curriculum innovation on learning outcomes, is based on systems thinking, an intellectual problem solving tool.

Systems Thinking

Systems thinking is a scientific analysis technique given prominence by Jay Forrester and his colleagues at the Massachusetts Institute of Technology. Work on computer modeling of systems thinking started over 30 years ago. Early models focused on urban growth and development and global patterns of the consumption of natural resources. 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.

The field of system dynamics provides a means to understand the behavior of complex phenomena over time. It is based on the concepts: (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 real-world systems over hypothetical time, are used to examine the structure of systems. Using simulation software, characteristics of selected variables can be altered and their effects on other variables and the entire system assessed. Thus, system dynamics 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).

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 (Structural Thinking Experimental Learning Laboratory with Animation; Richmond, 1985), has made it possible to operationalize these concepts on a microcomputer. STELLA capitalizes on the graphics and icon technology of the Macintosh microcomputer 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 the analytic and problem solving perspectives inherent in systems thinking. Hence, modeling now can be incorporated into science education at the secondary level. Recently the Mathematical Sciences Education Board (1987) recommended that modeling become a major emphasis in mathematics and science education (National Science Board Commission on Precollege Education in Mathematics, Science, and Technology, 1983). Because the costs of the Macintosh and STELLA now are affordable for many secondary schools, it is possible to implement the Board's recommendation on a wide-scale basis.

STACI

A research project designed to examine the cognitive and curricular impact of using systems thinking and STELLA in secondary school science and social studies courses is nearing completion. The project is a two-year research project conducted by Educational Testing Service under the auspices of the Educational Technology Center at the Harvard Graduate School of Education. STACI, which is now in its second year, examines the cognitive demands and consequences of learning from a systems thinking approach to instruction and from using simulation modeling software (see Mandinach & Thorpe, 1987a, 1987b).

The purpose of the study is to test the potentials and effects of using the systems approach in existing secondary school curricula to teach content-specific knowledge as well as general problem solving skills. The study also examines the effectiveness of using STELLA as a tool by which to teach systems thinking, content knowledge, and problem solving skills. The research focuses on (a) the learning outcomes and transfer that result from using such an approach and software in classroom settings, and (b) the organizational impact of the curriculum innovation.

Design

The study is being conducted at Brattleboro Union High School (BUHS), Brattleboro, Vermont. 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. Four teachers comprise the core of the systems group at BUHS. All were trained by experts to use the systems approach and integrated this perspective into their courses.

In the first year of the project, systems thinking was integrated into three general physical science, four biology, and three chemistry classes. An equivalent number of traditional (control) courses were taught concurrently by other members of the faculty. An experimental history course entitled War and Revolution also was taught with the systems approach. Table 1 presents the enrollment figures for the classes participating in the study during the 1986-1987 academic year.

Data Collection

Several types of instruments were used to assess outcomes in various stages of the research. These instruments included pretest, in-class, and posttest measures, which were used to assess ability, content-specific knowledge, systems thinking, and higher-order thinking skills (including general problem solving, metacognition, and self-regulation).

Pretests were used to assess subjects' ability, content-specific knowledge, and knowledge of systems thinking. Existing instruments were used or modified, and other tests developed where needed. BUHS supplied the students' most recent standardized achievement test scores. The California Achievement Tests (CAT) served as rough estimates of general ability. ETS also administered a small battery of tests, including the Advanced Progressive Matrices (APM) (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.

Modified versions of previous final examinations were administered to both the systems and traditional classes. The general physical science, biology (GPS), and chemistry 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. An initial assessment of systems thinking skills also was administered early in the semester to serve as a baseline for the experimental classes.

Teachers administered content-specific tests and exercises in their courses throughout the academic year. These examinations and exercises were roughly comparable in their subject-matter coverage. The teachers also prepared and gave common final examinations to their classes. Because traditional and systems thinking classes within a subject area received the same test, we were able to compare differences in content knowledge that resulted from using the systems approach.

ETS developed an instrument that was used to assess knowledge of systems thinking and STELLA. The instrument contained 76 items of increasing difficulty that measured a broad spectrum of skills along a continuum ranging from elementary concepts to complex modeling skills. Measures of systems thinking focused on concepts such as knowledge of graphing, equations, 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.

Results

The curricula. An integrative approach is used in the science courses, where the classes cover the same body of knowledge taught in traditional science curricula, but discussions of selected concepts and topics are supplemented with a systems thinking perspective. In these courses, students are learning concepts underlying model development and have an opportunity to experiment with existing models using STELLA. The modeling activities generally take two forms. Depending on the course, students are (a) required to develop their own models of scientific phenomena, or (b) are given existing models and are asked to alter particular parameters to examine the subsequent effects on the entire system.

These two distinct approaches to modeling are likely to produce different cognitive outcomes in terms of content knowledge and general problem solving skills. Although results are only preliminary at this time, there are indications that the manipulation of parameters in an existing model may promote scientific inquiry skills (e.g., understanding of causality and variation), and may directly influence the acquisition of content knowledge. In contrast, model building may be less explicitly related to the acquisition of content knowledge, yet may promote more general problem solving skills.

Teaching and learning outcomes. Although data collection in the project's first year was exploratory, several trends emerged from the instrumentation and observations. First, each of the science teachers adapted the systems approach to their courses in different ways, reflective of the particular content areas.

In GPS, the teacher introduced systems with graphing cause-and-effect relationships and simple arrow diagrams within the context of several topic areas (e.g., motion, magnetism). The concept of modeling was a recurrent theme throughout the course, with simple mathematical models developed to illustrate numbers or concepts.

Biology seemed to lend itself most readily to a systems approach. Because interrelationships among living systems is a key concept in biology, a number of relevant examples were identified to which systems thinking was applied. Students were introduced to the fundamental principles of systems, modeling, feedback, and causality. Modeling was integrated throughout the course, with presentations of models on topics such as oxygen production, metabolism, and population.

Chemistry was a more difficult course into which systems could be integrated. The teacher found the most appropriate chemical topics (e.g., relations among rates, time, and levels) and focused the approach there. Students were guided to develop systems models for chemical reactions. Structural diagrams and models provided instructional tools not only to illustrate the functions of and relationships among certain variables, but to hypothesize and then test changes in behavior of these variables over time, under different conditions. From these guided inquiry experiences, a generic understanding of the behavior of a set of interacting variables was developed.

These curricular differences made comparisons difficult, but enabled us to discern patterns of learning outcomes attributable to instructional foci. A developmental pattern emerged across the courses, with chemistry classes performing better than the biology students, followed by the GPS classes. The GPS teacher focused on measurement and graphing, concepts that underlie systems thinking at the most basic level. Consequently, these students were able to apply elementary systems principles, but were unable to develop models independently because they had not been sufficiently exposed to modeling. Parenthetically, observations of the GPS systems units used in the second year indicate that there is much greater emphasis on modeling and its practical application to physical phenomena. The teacher now is requiring students to develop elementary models of phenomena and solve problems using modeling.

In the first year, the biology teacher focused on model development, while the chemistry teacher emphasized the acquisition of content knowledge through the manipulation of modeling parameters. Student performance reflected these emphases. That is, the chemistry students were able to use modeling and the systems approach as problem solving tools for chemical and related problems. The biology students were adept at understanding the components of models, but were less able to apply the systems approach as a problem solving tool within the content area.

Results indicated that students were able to acquire knowledge of systems thinking concepts and apply them to scientific problems at varying levels of complexity and sophistication. However, because the curricula were not sufficiently developed, we are unable to make definitive statements at this time about the impact of the systems approach on the acquisition of content knowledge. A more thorough implementation of the curriculum currently is in progress. Thus, the second year testing will more systematically examine the impact of the curriculum differences on acquisition of content knowledge and general problem solving skills.

Systems Thinking Instrument (STI). The remainder of this paper will focus 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 from the first year administration then will be presented.

The STI consisted of 76 items that were intended to assess a range of skills thought to underlie systems thinking. A rational task analysis yielded 10 skills which then were made into subscales of varying lengths. Table 2 presents the alpha reliabilities for those scales. The 76 items on the total test yielded an alpha of .95, indicating that the test was extremely consistent across items and subscales.

The understanding of basic graphing concepts (e.g., labeling and scaling axes, coordinates) comprised the first subscale. Two other subscales focused on graphing skills. A first required the interpretation of graphs. That is, students were asked to interpret a graph and provide a verbal description. 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.

Two other scales focused on mathematical skills that relate to systems thinking. It was assumed that in order for students to become adept at using the STELLA software, they should have a basic understanding of how to solve simple equations. The second scale focused on students' understanding of graphical functions. Students were asked to write functions, find slopes, and graph functions.

Three scales were designed to assess knowledge of concepts that are critical in systems thinking. The first focused on variables. These items required students to differentiate between dependent and independent variables, place them on the appropriate axes, and interpret the graphs' meaning. Three items then targeted the notion of causality. Here students were asked to complete a causal relationship and a causal diagram. Defining causality leads directly into the concept of looping. Five items required that students either interpret or construct a causal loop diagram.

The final two scales measured skills and knowledge unique to systems thinking, STELLA, and modeling. The first consisted of simple identification items. One set of identification questions 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 identification items asked students to identify parts of a structural diagram (e.g., stock, connector, flow, converter). The final subscale focused on the construction and interpretation of systems models of varying complexity. 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. That is, 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, as well as for revisions of future instrumentation.

Second, and most importantly, we wanted the instrument to be reflective of and sensitive to differences in the systems curricula across the subject areas. As documented elsewhere (Mandinach & Thorpe, 1987c), 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. For example, the GPS classes stressed measurement concepts, but did not spend much time on modeling. The biology classes constructed some models, whereas the chemistry classes were given models and asked to modify them. The War and Revolution seminar achieved a high level of sophistication with systems thinking, modeling, and STELLA. 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. Analyses of performance in the project's second year will focus on the range of differences related to instructional emphases.

Table 3 presents the intercorrelations among the STI's 10 subscales. All correlations were found to be significant at the $p < .001$ level. Several relationships should be noted. First, the correlations among the graphing scales were quite strong. Correspondingly, interpretation and translation of graphs were related strongly and positively to the understanding of looping and causality. Performance on the two mathematically oriented scales (equations and graphing) were strongly related. Understanding the basic concepts of systems thinking (the Systems Identification scale) was strongly related to performance on the Graphing Translation, Causality, and Loops scales.

Interestingly, the Systems Identifications items were least related to performance on the Variables scale, which showed the lowest correlations with the other scales. As expected, the two systems scales showed the strongest relationship of all the intercorrelations on the instrument, $r(177) = .72$, p .001. Also as hypothesized, many of the subskills that underlie model construction were related to the Systems Interpretation scale.

Performance on both of the complex graphing scales (Interpretation, $r(177) = .51$, $p < .001$ and Translation, $r(177) = .56$, $p < .001$), Causality, $r(177) = .55$, $p < .001$, Loops, $r(177) = .61$, $p < .001$, and Mathematical Graphing, $r(177) = .61$, $p < .001$ correlated with students' ability to construct, interpret, and manipulate models.

Table 4 presents the descriptive statistics for the 10 subscales. These data are further broken down by course in Table 5. Although these tables present the raw scores, subscale- and item-level data discussed in the text are reported in terms of percent correct in order to aid interpretability.

On all three of the graphing subscales (Graphing, Graphing Interpretation, and Graphing Translation) developmental trends across courses were noted. Students in GPS generally scored the lowest, followed by biology, then chemistry. A ceiling effect was found for the War and Revolution students. They obtained a perfect score on the Graphing subscale and nearly perfect scores on Interpretation and Translation. In contrast, the GPS students obtained only 57, 35 and 34 percent correct on those scales.

More indepth analyses indicated particular problem areas experienced by the students. On the Graphing scale, 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.

Other weaknesses of the GPS students were noted on the Mathematical Equations and Mathematical Graphing scales. On the latter scale, these students showed no understanding of functions. They were unable to graph functions, determine slopes, or write a function when given a graphical representation. It is possible that freshmen-level mathematics courses had not covered these concepts. The GPS students achieved only 8 percent correct on this subscale, in contrast to 30 percent for biology, 53 percent for chemistry, and 100 percent for War and Revolution. The GPS students performed only slightly better on the Mathematical Equations scale ($M = .17$). They experienced substantial difficulty in solving simple mathematical equations. In contrast, the biology students achieved 67 percent correct, chemistry had 76 percent, and War and Revolution gained 98 percent correct.

A different pattern of performance was noted on the Variable subscale. Here, the GPS students ($M = .41$) outperformed those in biology ($M = .31$). Chemistry students achieved 58 percent correct, whereas War and Revolution obtained 86 percent. Examination of the biology students' responses indicated that they did not understand the difference between independent and dependent variables nor could they provide a rationale for how they defined the sets of variables. They also did not know how to graph the variables (i.e., on which axes were the independent variables and dependent variables to be placed).

Yet on the two scales more directly related to variation and causality (Causality and Loop), the developmental progression across classes reappeared. The items in the Causality scale required students to complete a causal statement (e.g., Amount of studying causes ______.), then make the causal loop. GPS students completed the statements, but had trouble 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. Completion of more complex causal loops were required within the Loop scale. Performance on this scale was slightly lower but analogous to the Causality scale ($\underline{M}_{GPS}$ = .27; $\underline{M}_{Biology}$ = .57; $\underline{M}_{Chemistry}$ = .62; and $\underline{M}_{W\&R}$ = .97). The GPS students experienced particular difficulty in translating a simple verbal description into a loop diagram. These results reflect the small amount of exposure the GPS students were given to loops.

On the scale that tested basic knowledge of systems thinking and STELLA, Systems Indentifications, there were several notable trends. Again there was a ceiling effect for the War and Revolution class ($\underline{M}$ = .94), particularly on the items that required identification of parts of a model (e.g., connectors, converters, flows, stocks). The GPS students ($\underline{M}$ = .42) had some trouble with the units of measure items as well as those that dealt with the parts of a model. Overall performance was roughly equivalent for the biology ($\underline{M}$ = .55) and chemistry ($\underline{M}$ = .59) classes. However, closer examination indicated that the chemistry students were more adept at the units of measure and variable items, whereas the biology students performed better on the structural modeling identifications. These performance trends were reflective of curricular emphases in the given courses.

A similar performance pattern was found for the Systems Thinking scale, where the focus was on modeling and interpretation of models. Given that the GPS students had little exposure to these complex concepts, they achieved only 19 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 ($\underline{M}$ = .33) and chemistry ($\underline{M}$ = .35) classes again performed similarly, whereas the War and Revolution seminar did extremely well ($\underline{M}$ = .86). These trends also were expected, given the structure of courses and how systems thinking was integrated into them.

What is striking about the Systems Thinking scale is that every item on the subscale exhibited the same performance trend across courses. These items emphasized various forms of representations (i.e., structural diagrams, verbal descriptions of problems or models) and required students to apply their knowledge of systems thinking in the form of model interpretation or construction. In all cases, the War and Revolution seminar either achieved or approached a perfect score. These scores were well above any of the science classes. The GPS students achieved the lowest scores, whereas the biology and chemistry classes performed similarly and between the two extremes.

Relationships among ability, achievement, content, and systems thinking tests. Analyses were conducted to see if general ability affected the pattern of correlations among content and systems thinking test performance, within subject area. Table 6 presents the first-order correlations for the chemistry classes. As expected, the pre- and posttests of chemistry were related, $r(54) = .33$, $p < .01$. Posttest performance also was related to the systems thinking instrument, $r(54) = .35$, $p < .01$. However, when partialling out the APM, no change in correlations was noted (Table 6). A different pattern emerged in the biology classes. The relationship between the biology pretest and performance on the systems thinking instrument was quite strong, $r(65) = .56$, $p < .001$. This magnitude of the correlation increased when APM scores were partialled out, $r(64) = .62$, $p < .001$. Clearly, general ability was a factor, to some degree, in the biology classes.

A third pattern was noted in the GPS classes, in which performance on the content pretest was not related to the systems instrument, $r(34) = .25$, ns. Yet the systems test was related to both the GPS posttest, $r(34) = .43$, $p < .01$, and to the APM, $r(34) = .47$, $p < .01$. When the APM was partialled out, the relationship between GPS and systems performance decreased, $r(33) = .32$, $p < .05$.

The relationship between the STI and CAT also were examined. These results are reported in Table 7. Achievement, as measured by the CAT, was related to performance on the STI. Not surprisingly, the reading subscale was most strongly related to the Graphing Interpretation ($r(119) = .52$, $p < .001$) and Translation ($r(119) = .51$, $p < .001$) subscales. These two scales required the most use of verbal skills. A similar trend was found for the language scale (Graphing Interpretation, $r(120) = .53$, $p < .001$; Graphing Translation, $r(120) = .51$, $p < .001$). Performance on the reading ($r(119) = .53$, $p < .001$) and language $r(122) = .56$, $p < .001$) was strongly related to the total STI score. Both the math-oriented subscales (Equations, $r(121) = .45$, $p < .001$, and Graphing, $r(121) = .53$, $p < .001$) showed strong, positive correlations with CAT mathematics performance. Additionally, Graphing Interpretation ($r(121) = .47$, $p < .001$) and Translation ($r(121) = .50$, $p < .001$) as well as total STI score ($r(121) = .52$, $p < .001$) were related to the CAT mathematics scale. Most of the STI subscales showed strong correlations with the CAT total score. Particularly strong were Graphing Interpretations, Graphing Translation, and Mathematical Graphing. The two total test scores were the most strongly related ($r(117) = .60$, $p < .001$) of all the correlations. Interestingly, the weakest relationship was between Systems Identifications across all of the CAT subscales.

Conclusions and Educational Implications

The data reported here are from the first year of the project. Thus, the focus is on the curriculum's impact on learning and teaching activities, rather than transfer. The impact of the curriculum innovation was apparent across the systems classes. Both teachers and students differed in the extent to which they used the systems approach effectively. Teachers carefully selected parts of their courses that were amenable to systems thinking. Extensive integration of systems and STELLA was accomplished in biology, the most natural fit for the approach. Systems thinking was integrated into GPS primarily in terms of measurement. Several complex systems models (e.g., pollution, reaction rates) were introduced in chemistry to illustrate the approach's applicability. Thus, students received different amounts and types of exposure to the approach across content areas.

Results indicated that students in the more advanced courses performed better on the systems thinking and computer-based activities. Chemistry students became most readily facile with the simulation-modeling software and drew direct applications to problems within the course. Biology students, who received the most extensive exposure to the systems approach, performed well when tested directly on the concepts and use of STELLA, but were less able to translate their knowledge and skills outside of biology to more general problems. This finding is most likely due to the teacher's emphasis on explicit examples within biology, not the applications of systems more generally. The physical science students' performance also reflected their teacher's emphasis on measurement concepts underlying systems thinking. These students performed well on measurement-related problems, but did less well on general problem solving or modeling-oriented exercises. Accordingly, these students were the least likely to generalize the systems skills beyond their course.

Also of importance are treatment differences reported elsewhere (Mandinach & Thorpe, 1987d). Differences in content performance emerged as courses became more advanced. There were no initial differences between systems and control classes in their knowledge of physical science, biology, or chemistry, as evidenced by tests administered in September. However, by June, results indicated that the systems approach was related to both qualitative and quantitative performance differences on specific items on the final exams. These items can be linked to curriculum modules taught with the systems approach. No differences were found between groups in physical science. Students in systems biology and chemistry performed better on items related to concepts taught with the systems approach. They also approached the problems in qualitatively different ways. Moreover, the systems students performed better on transfer tasks of general problem solving that related to skills and concepts underlying systems thinking.

The results raise questions about the applicability of the approach for different subjects and groups of students. We need to explore why the impact of the innovation was more apparent in biology and chemistry. Several explanations are possible. First and perhaps most likely was the type of exposure given to the physical science students and the naturalness of the approach's fit to the curriculum. A second hypothesis is that developmental differences mediated the results. Perhaps freshmen are not developmentally ready to profit from an instructional approach that emphasizes higher-order thinking skills. Another possbile explanation currently under examination is ability differences and subsequent self-selection into more advanced science courses.

Although the results from the project's first year are only preliminary, we have observed significant developments at BUHS as a result of the curriculum innovation. Many of the changes that we can report with confidence affect the teachers, curriculum, and the school as an organization. There is no doubt that the infusion of the systems thinking approach into the science classes has changed many of the teacher's instructional strategies and procedures by which to present traditional science 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. However, due to the developing nature of the curriculum, it is premature to discuss the learning outcomes data as anything more than exploratory at this time.

In conclusion, results from the project's first year indicated that the systems thinking approach and STELLA affected learning and teaching activites in the content areas as well as in general problem solving. The focus of next year's data collection will be to trace the differential performance of students who were exposed to systems and traditional classes as they take their next science courses. This will enable us to assess transfer across content areas as students become exposed to more courses taught with the curriculum innovation. In general, initial results are promising. The curriculum innovation appears to affect learning and teaching activities in positive ways. Further analyses and data collection will help to elucidate these findings.

References

Mandinach, E. B., & Thorpe, M. E. (1987a). Caveats and realities in technological curriculum innovation. Technology and Learning, 1(4), 1-3, 7, 9.

Mandinach, E. B., & Thorpe, M. E. (1987b). The systems thinking and curriculum innovation project. Technology and Learning, 1(2), 1, 10-11.

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

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

Mathematical Science Education Board. (1987, January). [Press release]. Washington, D.C.: National Research Council.

National Science Board Commission on Precollege Education in Mathematics, Science, and Technology. (1983). Educating Americans for the 21st century. Washington, D.C.: National Science Foundation.

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, N.H.: High Performance Systems.

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

Table 1

Classes and Enrollment Figures - 1986-1987

Class	Systems Thinking Classes	Systems Thinking Students	Traditional Classes	Traditional Students
General Physical Science	3	40(11)	3	57(3)
Biology				
Teacher 1	4	68(16)	2	32(12)
Teacher 2			2	29(10)
	4	68(16)	4	61(22)
Chemistry	3	57(3)	3	63(4)
TOTAL	10	165(30)	10	181(29)
War and Revolution	1	7(0)		

Note: The numbers in parentheses indicate the students for whom parental consent was either denied or could not be obtained.

Table 2

Reliabilities for the Systems Thinking Instrument Subscales

Subscale	Items	Alpha
Graphing	7	.51
Graph Interpretation	7	.72
Variables	4	.92
Graph Translation	7	.68
Causality	3	.92
Loops	5	.86
Math Equations	3	.67
Math Graphing	5	.87
Systems ID's	25	.92
Systems Interpretation	10	.87
Total Test	76	.95

Note. *n* = 179.

Table 3
Intercorrelations of Systems Thinking Instrument Subscales

	G	GI	V	GT	C	L	ME	MG	SI
Graphing									
Graph Interpretation	.57								
Variables	.39	.44							
Graph Translation	.53	.62	.40						
Causality	.40	.56	.33	.54					
Loops	.43	.64	.33	.59	.61				
Math Equations	.49	.51	.27	.46	.53	.42			
Math Graphing	.40	.44	.37	.53	.48	.42	.51		
Systems ID's	.32	.43	.28	.50	.54	.62	.38	.39	
Systems Interpretation	.34	.51	.32	.56	.55	.61	.42	.51	.72

Note. n = 179.
All correlations, $p < .001$.

Table 4
Descriptive Statistics for the Systems Thinking Instrument Scales

Subscale	Maximum Score	M	S.D.
Graphing	13	8.94	2.71
Graph Interpretation	20	10.99	5.10
Variables	12	5.30	3.84
Graph Translation	24	11.63	6.09
Causality	9	4.92	2.61
Loops	22	11.54	6.54
Math Equations	6	3.75	2.04
Math Graphing	8	2.75	3.21
Systems ID's	25	13.63	7.20
Systems Interpretation	46	14.90	10.76
Total Test	76	88.36	37.71

Note. n = 179.

Table 5

Course Breakdowns for the Systems Thinking Instrument Scales

Subscale	GPS		Biology		Chemistry		War & Rev.	
	M	S.D.	*M*	S.D.	*M*	S.D.	*M*	S.D.
Graphing	7.39	3.10	8.44	2.31	10.46	1.80	12.00	1.00
Graph Inter.	7.00	3.98	10.43	4.50	13.89	3.76	19.57	1.34
Variables	4.91	3.82	3.70	2.80	6.98	3.95	10.29	2.75
Graph Trans.	8.09	6.12	11.33	5.10	13.64	5.35	21.86	1.07
Causality	2.76	2.56	5.20	2.13	5.87	2.04	8.86	0.38
Loops	5.98	4.78	12.51	5.46	13.68	6.19	21.29	0.95
Math Eq.	2.04	2.36	4.00	1.52	4.57	1.51	5.86	0.36
Math Graph.	0.63	1.72	2.40	3.05	4.27	3.08	8.00	0.00
Systems ID's	10.39	6.52	13.81	7.09	14.82	6.84	23.57	1.62
Systems Inter.	8.96	7.79	15.27	9.51	16.23	9.84	39.71	6.90
Total Test	58.15	29.81	87.10	29.21	104.41	29.96	171.00	9.80

Note. n_{GPS} = 46. $n_{Biology}$ = 70. $n_{Chemistry}$ = 56. $n_{W\&R}$ = 7.

Table 6

First-Order and Partial Correlations Within Subject Area

	First-Order Correlations			Partialling Out APM	
	Pretest	Posttest	APM	Pretest	Posttest
Chemistry (n = 54)					
Systems	.18	.35**	.12	.18	.35**
Pretest		.33**	-.00		.33**
Posttest			.04		
Biology (n = 67)					
Systems	.56***	.40***	.43***	.62***	.39***
Pretest		.32**	-.00		.32**
Posttest			.11		
GPS (n = 36)					
Systems	.25	.43**	.47**	.24	.32*
Pretest		.60***	.08		.61***
Posttest			.36*		

Note. * $p < .05$. ** $p < .01$. *** $p < .001$.

Table 7

Correlations Between Achievement and Systems Performance

	Reading	Language	Mathematics	Total
Graphing	.40	.47	.41	.49
Graph Interpretation	.52	.53	.47	.59
Variables	.39	.45	.38	.46
Graph Translation	.51	.51	.50	.57
Causality	.38	.39	.42	.45
Loops	.34	.36	.33	.39
Math Equations	.23**	.29	.45	.35
Math Graphing	.47	.51	.53	.56
Systems ID's	.26**	.20*	.27	.27
Systems Interpretation	.36	.40	.31	.39
Total Score	.53	.56	.52	.60

Note. Chemistry and biology classes only.
* $p < .05$. ** $p < .01$. All other correlations, p .001.
$n_{reading} = 121$. $n_{language} = 122$. $n_{mathematics} = 123$.
$n_{total} = 119$.