DOCUMENT RESUME

ED 243 646 SE 044 371

AUTHOR Leinhardt, Gaea
TITLE Instructional Time: A Winged Chariot?
INSTITUTION Pittsburgh Univ., Pa. Learning Research and Development Center.
SPONS AGENCY National Inst. of Education (ED), Washington, DC.
REPORT NO LRDC-1984/8
PUB DATE 84
NOTE 42p.; Reprinted from C. W. Fisher, Ed. and D. C. Berliner, Ed., "Perspectives on Instructional Time". New York: Longman, Inc., (in press).
PUB TYPE Journal Articles (080) -- Reports - Research/Technical (143)

EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Academic Achievement; *Achievement Gains; Elementary Education; Elementary School Mathematics; *Elementary School Science; Learning Disabilities; *Mathematics Achievement; Mathematics Instruction; *Reading Achievement; Reading Difficulties; Science Instruction; *Time Factors (Learning); *Time Management; Time on Task

ABSTRACT

Overviews of three studies which focused on achievement as related to instructional time are presented. The first study investigated whether or not time spent in science with fourth and fifth graders (N=86) detracted from learning in more basic areas. The second study examined reading and mathematics achievement in grades 1 and 3 as related to initial performance and to four classroom processes. These processes were represented by four constructs: opportunity, motivators, instructional events, and structure. Opportunity consisted of two variables, time and curriculum overlap, with time being estimated by attendance, allocation, task rate, enrollment, and transfers. The third study investigated the nature of reading difficulties in learning disabled (LD) classes, type of student activities which lead to greatest improvement in reading test performance, and what types of instructional situations generate these student activities. Among the basic findings which have emerged from these studies are those indicating that time is overlapping and not mutually exclusive within a student or teacher, even when the focus is primarily on one type of activity; that time on task is not the same as time on the right task; and that allocated time may not be the upper bound for engaged time. (JN)

LEARNING RESEARCH AND DEVELOPMENT CENTER

INSTRUCTIONAL TIME: A WINGED CHARIOT?

1984/8

GAEA LEINHARDT

INSTRUCTIONAL TIME: A WINGED CHARIOT?

Gaea Leinhardt

Learning Research and Development Center
University of Pittsburgh

1984

Reprinted by permission from C. W. Fisher & D. C. Berliner (Eds.), *Perspectives on instructional time.* New York: Longman, Inc., in press.

The research reported herein was supported by the Learning Research and Development Center, funded in part by the National Institute of Education (NIE), U. S. Department of Education. The opinions expressed do not necessarily reflect the position or policy of NIE, and no official endorsement should be inferred.

INSTRUCTIONAL TIME: A WINGED CHARIOT?

Gaea Leinhardt

Learning Research and Development Center

University of Pittsburgh

May, 1981

But at my back I always hear
Time's winged chariot hurrying near;
And yonder all before us lie
Deserts of vast eternity.
Andrew Marville

As I understand our task, we were to present a brief overview of our own research and findings, summarize the major points that we have learned about time, and give some estimates as to what we think the future of this line of research holds. I would like to report on three fairly separate studies that I think are illustrative of the way the field has gone, and certainly the way my own thinking has gone.

Science Study

The first study that I'd like to describe is The Science Study. This study was carried out in 1974 as part of the evaluation of the Learning Research and Development Center's (LRDC) Individualized Science Curriculum (IS) (Champagne & Klopfer, 1974). In evaluating IS we were faced with the difficulties of having a weak set of dependent measures, and of having no point of comparison. There are virtually no norms for sucess or failure in elementary science instruction (Leinhardt, 1977). In general, in the elementary schools in Western Pennsylvania, students are not exposed to science education, so that any child that is taught anything is clearly going to look better than a child that hasn't been taught anything at all. One serious concern, even in 1974, was whether or not time spent in science detracted from learning in the more basic areas. What we did was to construct a composite pre and posttest battery and estimate time spent in science, math, and reading instruction. The pretest consisted of the previous year's end-of-year SAT (Stanford Achievement Test; Madden, Gardner, Rudman, Karlsen, & Merwin, 1972) stanines in science, math, and reading; the posttest consisted of the Spring 1975 stanines. The modified allocated time measures consisted of students' own punched clock records in math and reading and teacher logs of time spent in science. All the time estimates are closer to allocated than engaged estimates, but they are lower than strict schedule-based allocations. The study was carried out in Oakleaf School, a demonstration school for LRDC and included all 86 fourth and fifth graders enrolled. The results indicated that after pretest, the only other important predictor of test performance was time spent in science.

Insert Table 1 here

Table 1 shows the means, standard deviations, correlation matrix, partial correlation matrix, and regressions. Notice that the zero-order correlation of time spent in science with outcomes is negligible, whereas the math time is quite substantial. However, when the partial correlations are examined, science time emerges as an important predictor. The regression results are similar. Why is this? Well, first of all, science time is not correlated with initial ability, whereas, time spent in math and reading is. Secondly, time spent in science reinforces reading comprehension by having action-based (follow the directions) short passages (two to three pages). Third, the science material is very well written. There is also some limited amount of measuring and calculation that is required which reinforces math skills. The initial analysis consisted of correlations and partial correlations. By 1974, we knew that the initial ability of the student was the most powerful predictor of outcome and that many of one's favorite classroom process variables, time among them, was confounded with initial ability. We knew, therefore, that simple correlations with outcome might be deceiving because one was really capturing only relationships between process and input, so we knew to partial them. The regression analysis has been added in an attempt to have comparable results across the studies. The regression results are consistent with the partial correlations.

One value of this study, it seems to me, is to point out the need for caution. As we strive to get schools and teachers to increase the engaged time spent by their students in academically fruitful activities, we must be careful about the inappropriate increase or decrease in allocated time in subject areas. As allocated time in language arts creeps up from 60 to 120 minutes a day, time is taken away from substantive areas such as science and social studies, areas which undoubtedly have benefit for their own sake but which may have unanticipated benefits for the so-called basics as well.

IDS

Let me turn now to the second study. This study is the Instructional Dimensions Study (IDS). It was designed by William Cooley and myself at the request of Congress and NIE. It was carried out by Kirschner Associates and Education Turnkey. Hugh Poynor, Lee Poynor, Charles Blaschke, and Jack Sweeney were the primary movers in the actual conduct of the study, and William Cooley and I wrote up the results and reported on them (Cooley & Leinhardt, 1980).

> The design of IDS had several distinguishing features, including its use of a model of classroom processes, the direct measurement of these processes, and distinctive outcome measures, sampling procedures and methods of analysis. Each of these features is briefly described . . .
>
> A model of classroom process. The model used by IDS for the study of classroom processes is illustrated in Figure 1 (Cooley & Leinhardt, 1975c; Cooley & Lohnes, 1976). The figure identifies constructs--sets of variables--that we believe are necessary to explain the variation in student performance that occurs among classrooms after an extended period of instruction in those classrooms. It suggests that criterion performance is a function of initial student performance and of certain classroom processes that occur in the interval between the assessment of initial performance and the assessment of criterion performance (it does not, however, specify the exact nature of those functional relationships).

Classroom processes are represented by four constructs: opportunity, motivators, instructional events, and structure.

Insert Figure 1 here

The opportunity construct represents the student's opportunity to learn what is tested in the criterion performance measures. This construct incorporates how time is spent in classrooms and the similarity of the curriculum to the tests. The motivators construct includes those aspects of the curriculum and in-class interpersonal behavior that encourage learning. Instructional events include the content, frequency, quality and duration of instructional interactions. Structure, the fourth construct, considers the level of organization of the curriculum, the specificity of the objectives, and the manner in which a student and a curriculum are matched. . .

<u>Measures of classroom processes</u>. From published research we identified the measurable features of each construct and selected measurement procedures that minimized classroom disruption and maximized precision. Measures within a construct were combined into variables, which were in turn combined into sets of varibles representing the four constructs. . .

Measures of outcome. The study used commercially available achievement tests in reading and mathematics. . . The test battery selected was the Comprehensive Test of Basic Skills (CTBS, 1974) of the California Test Bureau; level B was used in grade 1, and level 1 in grade 3. . .

Sampling procedures. The sampling scheme emphasized the need to achieve variance in classroom processes, rather than representativeness of current compensatory practices. This variation in processes was obtained by using descriptive information about the programs in use in a large number of school districts and then selecting classrooms from those districts that would increase the likelihood of process variation. . .

Methods of analysis. The analytic procedures recognized that various classroom practices naturally occur in combinations that are not susceptible to experimental control. At this stage of our understanding we need a technique for determining the relative importance of different practices in explaining variations in student achievement. (Cooley & Leinhardt, 1980, pp. 8-9.)

The analysis that was used in IDS was commonality analysis. The results strongly suggest that the most important of the predictors in the model were pretest and opportunity to learn. Opportunity consisted of two variables: time and curriculum overlap. Time was estimated by attendance, allocation, and on task rate as well as enrollment and transfers; overlap used a teacher-based estimate and a curriculum

6

nalysis estimate.

In first grade, for both reading and math, the pretest and the opportunity variables were the most useful predictors of gain in achievement; opportunity was by far the best predictor of reading gain; whereas pretest and opportunity had identical uniquenesses in math.

Third-grade reading showed the highest correlation between pretest and posttest (r =.86), leaving the smallest amount of reliable variance for other predictors to explain. Even so, the usefulness of pretest and the opportunity in predicting gain is about as great as for first-grade mathematics. . . Third-grade math also shows the importance of the opportunity construct; it was, in fact, the only consruct that had a significant uniqueness for this sample (in all other samples, at least two constructs were significant). In summary, the most useful construct for explaining achievement gain is the opportunity that the children had to learn the skills assessed in the achievement te*. . . (Cooley & Leinhardt, 1980, p. 20.)

<u>The importance of curriculum content</u>. Curriculum content is one area in which classroom practice seems to make a difference. Instructional techniques, both curricular and interpersonal, do not compensate for missing content. Children perform better on a test if they have been exposed to both the <u>form</u> of the test items and the <u>content</u> covered by the test. This point may seem trivial, but it is in fact quite

> important. Both what is taught and what is tested represent samplings from domains. Curriculum and test can clearly emphasize different content. The results of IDS show that students are most likely to answer correctly if they have been taught the specific material covered by a test and if they have been frequently exposed to the test format. (Cooley & Leinhardt, 1980, p. 22.)
>
> Of the four process constructs assessed in IDS, opportunity is the strongest, most consistent predictor of achievement gains. Within opportunity, the variable that stands out is that which assesses the degree to which the children in the classroom had an opportunity to learn what was in the end-of-year achievement test.
>
> . . . .
>
> One implication of the importance of opportunity is that evaluation studies that do not include information on how time is used or on the degree to which the curriculum included what the achievement test measured are suffering from specification error. That is, there is a danger of attributing instructional effectiveness to specific programs or ways of teaching when it is really a matter of the curriculum content being a good fit to the particular achievement test that happened to be selected. (Cooley & Leinhardt, 1980, p. 23.)

What IDS helps us to see is that until we control for the massive variation by classroom in time spent in academic areas and the content covered, we are unlikely to find other process variables that are important. That does not mean that these other variables are not important, only that discovery of their importance is hampered.

<u>LD Study</u>

Let me turn now to the third piece of research on which that I want to report. This study was carried out between 1977 and 1980. It is a study of reading instruction in classrooms for the learning-disabled. It was carried out by William Cooley, Naomi Zigmond, and myself (Leinhardt, Zigmond, Cooley, 1981).

> Reading instruction and its effects were examined for 105 students in elementary classrooms for the learning-disabled. Extensive detailed observations of students, teachers, and instructional material were used to explore the plausibility of a causal model of the effects of reading behaviors and instruction on students' reading performance. The results indicate that 72 percent of the variance in posttest reading scores can be explained by a model that includes a pretest, three student reading behaviors, and instructional overlap; and that 59 percent of the variance in student time spent in reading can be explained by a model that includes pretest, teacher instructional behaviors, teacher affective behaviors, and instructional pacing. (Leinhardt, Zigmond, & Cooley, 1981, p. 1.)

. . . .

The number of children classified as learning disabled has risen dramatically during the past decade, from 120,000 in 1968 to 1,281,395 in 1979. One of the primary reasons that elementary school children are diagnosed as learning disabled and assigned to special classrooms is that they exhibit relatively poor performance in reading. Within LD classrooms, reading instruction is given considerable emphasis, and a wide variety of instructional practices and materials are tried. . .

In this investigation, we pursued answers to three major questions: (a) What is the nature of reading activities in LD classes? (b) What types of student activities lead to greatest improvement in reading test performance? and, (c) What types of instructional situations generate these student activities?

Two basic assumptions about effective reading instruction guided our data collection and analysis activities. First, we assumed that what students learn is a function of what they do, and that features of the curriculum and teacher behaviors influence what students do rather than directly influencing what they learn. Second, we assumed that beginning reading activities fall into three broad categories: those directly related to the reading task in that they involve the student responding to print; those that indirectly support some aspect of reading, but are not reading (e. g., listening to the teacher, or talking about a story); and those that are so tangential to the acquisition of reading competence as to be non-reading (e. g., working with perceptual training boards, or doing auditory discrimination tasks). We imposed this view of reading instruction and reading behaviors on our observational system; we did not define reading simply as what goes on during allocated reading time, nor as what a test manual describes as components of reading. (Leinhardt et al., 1981, pp. 3-4.)

. . . .

Data on the classrooms, students, and teachers were obtained in a number of ways. Teachers were observed directly and interviewed. Students were observed directly, and their work products and assignments monitored. In addition, students were tested and their prior test data were also recorded. For all data collected the individual child was the unit of observation and analysis. This was justified because of the totally individualized instructional programs. Classroom observations and teacher interviews yielded information on a wide variety of variables. (Leinhardt et al., 1981, p. 5.)

. . . .

Both the teacher and student were observed directly in this study. Several measures of their behaviors were recorded using a time sampling procedure (Student-Level Observation of Beginning Reading) (SOBR) . . . (Leinhardt & Seewald, in press).

Teacher behaviors. Teacher behaviors were observed and linked to the student(s) to whom they were directed. The two basic areas of teacher behavior recorded were instructional and affective. Instructional behaviors included model presentations, explanations, feedback, cueing, and monitoring. these measures were obtained from time samples taken every five minutes for one hour. . . These separate measures were combined into a single estimate of the time an individual child received teacher instruction each day. Affective behaviors included the number of reinforcers received by each child per day and the cognitive press exerted by the teacher toward a child. . .

Student behaviors. Figure 2 displays the schema used to organize the observations of student behavior. The system used in this study (SOBR) was designed to assess what students actually did during reading instruction. A detailed analysis of reading behaviors coupled with an assessment of the tasks presented to children during reading formed the bases of the observational system. The system was intended to be exhaustive, in that all of the students' time during observation would be accounted for. Students could be reading or not. If they were reading, these behaviors were classified further as direct or indirect. Direct reading behaviors included oral and silent reading of letters, words, sentences, and paragraphs. Indirect reading behaviors were those activities that were assumed to be related to reading but in which the student was not engaged in responses to print in the "normal" way. By normal we mean going from print (as stimulus) to sound or silent reading (as response) as opposed to finding letters with a specific sound. The indirect reading behaviors include story discussion, circling pictures with a common phonetic element, listening, and writing, whether copying or spelling. The combined measures of reading used in this study included the amount of time per day a child was reading aloud, reading silently, or engaged in indirect reading behaviors.

Insert Figure 2 here

In addition to reading categories there were five non-reading categories observed and recorded. The non-reading categories were: waiting for something or someone; academic activities other than reading; management activities including preparing for a task or wrapping-up after a task is completed; absent from school; or, out of the room at the time of observation. Further, off-task during reading was distinguished from off-task during other kinds of activities. . .

The classroom observations of the 105 LD students took place over a 20-week period, from December 1978 to May 1979. For each classroom, we randomly sampled twenty one-hour observation sessions in the morning and ten in the afternoon. Thus, each classroom was observed for about 30 hours over the 20 weeks. This considerable effort at extensive sampling was to assure adequate coverage and stability of the phenomena of interest (Karweit & Slavin, 1980).

Raw counts were weighted so that from our thirty observations per student we could estimate the classroom behaviors of each student between pretest and posttest. Weights were a function of the reciprocals of the sampling ratios, that is, if we observed a student for 20 morning hours, and s/he was in school for a total of 320 morning hours, then the weight for the AM observations would be 320/20, or 16. Similarly, if a student was observed for 10 PM hours, and s/he was in school for 300 afternoon hours, the estimates for the PM strata would be weighted by 30. The weighted estimates were then scaled so that they approximated the minutes per day a child averaged in a particular activity. The complete sampling and weighting procedures are described in Cooley and Mao (1980). (Leinhardt et al., 1981, pp. 10-14)

. . . .

The major objectives of this study were to determine how the activities that students engaged in affected their test performance, and how a variety of instructional features influenced student behavior. These relationships were examined by a series of multiple regressions. The variables to be included in each regression were determined by a model for explaining reading achievement. . .

Figure [3] displays the combined causal model of how the variables are assumed to be influencing each other in the classroom. It summarizes the two functions described above. Solid black lines indicate significant relationships that are time driven, and in which the directionality of the arrow seems clear (e.g., pretest to posttest). Broken black lines indicate significant relationships in which we assume a causal directionality, but in which both variables were measured at approximately the same time (e.g., teacher instructional behaviors and time in reading activities). Dotted lines indicate relationships that we predicted would be significant but were not.

Insert Figure 3 here

It is important to note that in some cases combined sets of reading variables are used and in some cases single reading variables are used. For example, teacher affective support is aimed at increasing all of the students' reading activities so the total reading time was the appropriate dependent variable. Posttest, however, is assumed to be influenced primarily by silent reading, somewhat by oral reading, and very little by indirect reading activities, so these are examined separately in their role as explanatory variables. This distinction is not made merely for the purpose of conserving degrees of freedom, but because of the logic of the relationships.

The main point of Figure [3] is to show that posttest is assumed to be dependent on student behaviors and instructional content; student learning behaviors are assumed to be influenced by prior test performance and teacher behaviors. . . The regressions reported below the figure indicate that posttest . . . performance is significantly influenced by pretest, silent reading time, and overlap, but not significantly influenced by oral reading or indirect activities. (Oral reading time was significant when an oral reading test was used and no overlap estimate was included.) These results suggest that an average of one minute per day of additional silent reading time increases posttest performance by one point. An increase of five minutes per day would be equivalent to about one month (on a grade-equivalent scale) of additional reading achievement.

Turning to the factors that increase time spent by students in direct and indirect reading behaviors . . . From Figure [3] we can see that total time in reading was expected to be influenced by what students knew in the beginning, what the teacher taught them, how the teacher encouraged or cajoled them, and the pacing of instruction. All but the last of these are verified. The regression suggests that an increase of one minute of teacher instruction per day gains a minute of student reading time, and each reinforcer also increases daily reading time. Given these results, teachers could increase instructional time to several students at once. The pacing result, that is, lack of a significant relationship, may be explained by the fact that these teachers tended to adjust the pacing to the level of student ability so that students of a given ability were not differentially "pushed" by longer or shorter assignments. Indeed, the strong zero order relationship, .46, between pacing and pretest suggests this, but its partial regression is not significantly different from zero. (While this is different from the results of Barr, 1980, the model being tested is also different.)

We spent over two years studying these classrooms. Much of the information we obtained is captured by the specific variables and analyses . . . presented so far. Much more, however, was also learned. In fact, after observing classrooms and teachers (some for more than 70 hours), [I] feel obliged to report what we have learned from both our more formal and less formal analyses. It is the latter experience that grounds the former and makes it more than a collection of numbers. Thus, there are two classes of findings from this study: those that deal with the state of the art of research on effective classroom processes and those that deal with the substantive import of our findings.

After decades of effort, research on instructional domains has reached a point where classroom processes can be measured with reliability and validity. This is due to the: improved strategies for sampling the instructional domain; precision with which observational measures can be crafted; incorporation of interpretable metrics into the observation system (namely time); and, heuristic value of a causal scheme

> Improved strategies for sampling include capturing more of the time than most previous work (30 hours), and randomly sampling the time of observation. Previous work has used as little as two hours and has intentionally waited until instruction "began", thus inflating estimates of instruction. In this study, the use of schema to organize the observational measures dramatically improved the reliability of observers and eased final analysis. Using a unifying metric for the majority of measures improved analysis and interpretability as it has for other researchers in the area, notably the BTES work. Using a causal model has helped to clarify the paths through which variables may be operating. Specifically, it is student behaviors during instruction that influence student learning, while teacher behaviors influence student behaviors.
>
> (Leinhardt et al., 1981, pp. 17-23.)

This set of three studies traces both thinking and findings about instructional time. In the Science Study, we considered a modified estimate of allocated time to be useful in understanding student end-of-year performance. At that stage in my thinking, time was simply a blocked competitive resource that needed justification. In IDS, somewhat more complex measures of time, attendance, and on-task rates were used to try to get at something more relevant for instruction than allocation. As we began to understand the importance of content covered, we tried to move toward notions of time spent in appropriate tasks. The L-D Study carefully defined the instructional domain, heavily sampled it and scaled it in terms of time.

Having reviewed this research, I would like to turn now to six basic findings that have emerged from the study of instructional time.

1. First, time is most usefully thought of as a metric not a variable.

2. Second, time spent is overlapping and not mutually exclusive with one organism even when the focus is primarily on one type of activity.

3. Third, observations and interpretations must be at the same level. This is not the same as unit of analysis. It is the problem of kids spending four to ten minutes a day reading while teachers are spending two to four hours a day teaching reading.

4. Fourth, allocated time is not the upper bound for engaged time in some situations, while it is the upper bound for others.

5. Fifth, time on task is not the same as time on the right task.

6. Sixth, to obtain a stable, generalizable estimate of students' time usage, there must be extensive sampling and extremely well-defined variables.

Let us now look at each of these points separately.

Time is a metric not a variable. By that I mean the passage of time is in itself of no particular value to any research effort or any question. It is time as a metaphor or time as a metric for something else. As a metaphor, it is useful in considering time as a resource, a relatively finite resource with respect to the consumers although not particularly finite with respect to itself. As a metric, it is useful because it permits us to scale variables in a way that permits their addition and subtraction. However, caution is needed once behaviors or variables have been scaled in the metric of time. They appear to have an equivalence that may be false. One minute of teacher contact in Category Y is not the same as one minute of student behavior in Category X, and issues of intensity and density need to be addressed. Time can be a deceiving metric, but it is useful with respect to policy recommendation and with respect to notions of improvement and change. It is also the metric that gives us the greatest shock value. Remembering that we are only estimating from samples, we seem to discover surprisingly small amounts of time spent in activities that we presume have high payoff for a variety of academic areas.

<u>Time is overlapping and not mutually exclusive</u>. By this, I am referring to the fact that several behaviors go on at once within one child or teacher. Two or more things can occur simultaneously, so that they slightly overlap, or exclusively. This indicates that the summation of time must be done with care and precision. We must know the boundaries at which we are summing. For example, consider a table that displays the following percentages, "80 percent of the time reading, 75 percent of the time in social contact, 65 percent of the time in contact with the instructional leader." When all these percentages are above 50, the interpretability gained by using time is lost. Caution is needed in deciding which element can be carved neatly away from which other element.

Observations and interpretations must be at the same level. My third point is the notion that different actors in the same environment, spend very different amounts of time in the same activity and that estimating one from the other is difficult. If we study the time use of children and attach teacher behaviors to the child's time, we cannot estimate the teacher's use of time. Because a child receives less than a minute a day of cognitive instruction and about 15 minutes a day of instruction relevant in some way to reading, we cannot infer that the teacher wasn't working at teaching reading. Teacher time has to be estimated separately, and linkage may be a problem. Allocated time may not be the upper bound for engaged time. Clearly, allocated time is pretty close to being the upper bound for mathematics instruction. Very little formal mathematics instruction goes on in any class other than math class. My concern with this notion of allocated time being the upper bound for engaged time stems from the Science Study described earlier. Remember, time spent in science influenced science, reading, and math achievement. This is not simply a matter of redefining allocation. It is a problem in understanding the total instructional day. We need that understanding in order to address the question, "Should we allocate more reading time in order to get more engaged reading time, or should we permit allocated reading time to stay as it is?" "Buying" more reading allocation will occur at the cost of social studies or science or music, however, that time spent will also have payoff in the area known as reading comprehension. The policy implications of this issue are clear.

Time on task is not the same as time on the right task. My fifth point is one that we certainly all understand, but needs to be repeated frequently. Time on task is synonymous with engagement in relevant tasks only when the task content has been controlled for. Much of the time on task literature merely takes children being on task during the time allocated to Instructional Area X with no regard to the content of the task they are performing. This notion fits very closely into the idea that time is a metric. In order for one to do a study of time that is meaningful, a detailed analysis of the content of instruction must be carried out. Simply taking a percentage of on-task behavior and multiplying it times allocation will not give you an engagement rate that is useful, as we learned in IDS.

There is a need for extensive sampling and well-defined variables. The sixth and final point deals with methodology. In a generalizability study that was carried out as a part of the L-D Study, we discovered that in order for estimates of student behavior to stabilize, we needed to have extensive sampling of those behaviors (Lomax, 1980; Cooley & Mao, 1980). Specifically, 20 hours were needed as opposed to the two or three usually recommended. This large time sampling was not necessary for all variables, but it was necessary for many of the less frequently occurring ones. In addition to sampling the behaviors, adequately defining them carefully is as important. As we get better at defining tasks, activities, and behaviors, our ability to connect them becomes improved.

Future.

Having reviewed samples of the last decade's work, I would like now to turn to what we need to do in the future. One area that I think would be interesting would be to unravel teacher time. We have done a fairly good job of understanding a student record; that is, taking a child, explaining how that child spends a day, and tagging on to it the resource that that child has access to, namely the amount of teacher time the child has access to. But we haven't looked at it from the other end; that is, the understanding of how teachers distribute their day in terms of time. I think that we're going to need to understand that if we're going to reform how teachers get children to spend time. The best work that I know of that's going on now is the teacher planning and teacher decision making studies. Linking that work with the time issue would be exciting.

The second area that I think we ought to work on is the notion of multiple payoffs, that is, the notions that were initially described in the Science Study. Clearly, we begin to get payoff especially in reading for things like time spent in drama, time spent in areas that involve high content and contact with print but that aren't called reading. But we're going to find out about the efficacy of using that time only as we begin to analyze the specific skills that are required for the test performance with criterion measures or whatever you want to use, and we do a similar analysis of the other curriculum content areas.

In the future, we need to analyze carefully the task components in a variety of subject matter domains; science, reading, social studies, mathematics, and music would be my candidates. These analyses then need to be compared so that we can construct portraits of a student's day and map those on to the more complex aspects of student achievement that we all hope to use someday. Figure 4 displays a simple version of what I have in mind. The point is to examine how various areas of instruction support each other in meaningful ways. Humanistic educators have long argued against back to basics, in part because of the enrichment value that is lost when the other subjects are deleted. But more may be being lost. There may, in fact, be some specific "basic" skills being lost as well. For example, in my purely hypothetical case, reading connected prose silently takes up 15 minutes in reading, 10 minutes in math, and 6 minutes in science. Oral reading takes up 10 minutes of reading, 3 minutes of math, and is negligible in science. Writing is done for 5 minutes of reading, but 20 minutes of science time. If we remember why we are supposed to study the "basics", perhaps it will be no surprise that children in subject area classes spend time exercising these basic skills and engaging in them. In fact, in at least one school in Pittsburgh, there is no allocated reading time after third grade--making way for expanded social studies, civics, and science classes. Mapping these activities out and demonstrating their relationship to outcomes is a non-trivial task and will take us quite a few person years of work. Simultaneously, we have to build more complex analyses of existing tests and hopefully track areas of instruction where time is spent but never assessed (clear oral presentation, for example).

Conclusions

In sum, my recommendation is that we continue to investigate the pressing problems of education; how to help poor achievers learn more easily and in greater depth; how to keep the gifted and talented challenged; how to make our schools nurturant, attractive places in which children experience the beginnings of the wondrous adventure of life; and that we continue these investigations using time as the powerful metric we have seen it is but while wending our way carefully between the Scylla and Charybdis of overly simplistic interpretations and overly complex obfuscations of reality.

References

Anderson, L. W. Time and school learning. Abstract unpublished doctoral dissertation, University of Chicago, 1973.

Anderson, L. W. New directions for research on instruction and time-on-task. Paper presented at the annual meeting of the American Educational Research Association, Boston, April 1980.

Anderson, L. W., & Scott, C. C. The Classroom Process Scale (CPS): An approach to the measurement of teaching effectiveness. Paper presented at the annual meeting of the American Educational Research Association, Toronto, March 1978.

Anderson, L. W., Scott, C. C., & Reynolds, A. Improving classroom instruction through an analysis of student errors. Paper presented at the annual meeting of the American Educational Research Association, Toronto, March 1978.

Barr, R. School, class, group and pace effects on learning. Paper presented at the annual meeting of the American Educational Research Association, Boston, April 1980.

Block, J. H. The effects of various levels of performance on selected cognitive, affective, and time variables. Unpublished doctoral dissertation, University of Chicago, 1970.

Bloom, B. Time and learning. American Psychologist, 1974, September, 682-688.

Brown, B. W., & Saks, D. H. The allocation and use of instructional time in reading and mathematics: Some theory and evidence. Paper presented at the annual meeting of the American Educational Research Association, Boston, April 1980.

Champagne, A. B., & Klopfer, L. E. An individualized elementary school science program. Pittsburgh, PA: University of Pittsburgh, Learning Research and Development Center, 1974. (LRDC Publication 1974/10)

Comprehensive Tests of Basic Skills. Monterey, CA: CTB/McGraw-Hill, 1974.

Cooley, W. W. Explanatory observational studies. Educational Researcher, 1978, 7(9), 9-15.

Cooley, W. W., & Leinhardt, G. Compensatory education study: Design for a study of the effectiveness of individualized instruction in the teaching of reading and mathematics. Pittsburgh, PA: University of Pittsburgh, Learning Research and Development Center, 1975. (a)

Cooley, W. W., & Leinhardt, G. Design for the individualized instruction study: A study of the effectiveness of individualized instruction in the teaching of reading and mathematics in compensatory education programs. Pittsburgh, PA: University of Pittsburgh, Learning Research and Development Center, 1975. (b)

Cooley, W. W., & Leinhardt, G. _The application of a model for investigating classroom processes._ Pittsburgh, PA: University of Pittsburgh, Learning Research and Development Center, 1975. (LRDC Publication 1975/24) (c)

Cooley, W. W., & Leinhardt, G. The Instructional Dimensions Study. _Educational Evaluation and Policy Analysis_, 1980, _2_(1), 7-25.

Cooley, W. W., & Lohnes, P. R. _Evaluation research in education._ New York: Irvington Publishers, 1976.

Cooley, W. W., & Mao, B. J. _The sample of classroom time observed._ Paper presented at the annual meeting of the American Educational Research Association, Boston, April 1980.

Corno, L. _Classroom instruction and the matter of time._ Chapter 8. Unpublished manuscript, n.d.

Derevensky, J. L., Hart, S., & Farrell, M. _An examination of achievement related behavior of high and low achieving inner city pupils._ Paper presented at the annual meeting of the American Educational Research Association, Los Angeles, April 1981.

Devine, V. T., & Tomlinson, J. R. The "Workclock": An alternative to token economies in the management of classroom behaviors. _Psychology in the Schools_, 1976, _13_,(2), 163-170.

Ebert, M. K. _An investigation of the causal relationships among locus of control, environmental motivators, engaged reading time, and reading ability._ Unpublished doctoral dissertation. Pittsburgh, PA: University of Pittsburgh, Learning Research and Development Center, 1980.

Fisher, C. W. Academic learning time as instructional feedback. Paper presented at the annual meeting of the American Educational Research Association, Boston, April 1980.

Fisher, C. W., Filby, N. N., Marliave, R., Cahen, L. S., Dishaw, M. M., Moore, J. E., & Berliner, D. C. Teacher behaviors, academic learning time and student achievement: Final report of Phase III-B, Beginning Teacher Evaluation Study. San Francisco: Far West Laboratory for Educational Research and Development, 1978.

Frederick, W. C., & Walberg, H. J. Learning as a function of time. Journal of Educational Research, 1980, 73, 183-194.

Harnischfeger, A., & Wiley, D. E. Time allocations in fifth grade reading. Paper presented at the annual meeting of the American Educational Research Association, New York, April 1977.

Harnischfeger, A., & Wiley, D. E. Teacher resource allocation: Consequences for pupils. Chicago: ML-Group for Policy Studies in Education, CEMREL, Inc., 1978.

Hess, R. D., & Takanishi, R. The relationship of teacher behavior and school characteristics to student engagement. Technical Report No. 42. Stanford, CA: Stanford University, Stanford Center for Research and Development in Teaching, 1974.

Israelite, L., & Deen, A. The effects of self and teacher evaluation on student performance and continuing motivation. Paper presented at the annual meeting of the American Educatinal Research Association, Los Angeles, April 1981.

Karweit, N. A reanalysis of the effect of quantity of schooling on achievement. *Sociology of Education*, 1976, 49(July), 236-246.

Karweit, N. *The Organization of time in schools: Time scales and learning.* Unpublished manuscript, Johns Hopkins University, Center for Social Organization of Schools, 1977.

Karweit, N. L., Fennessey, J., & Daiger, D. C. *Examining the credibility of offsetting contextual effects.* Report No. 250. Baltimore, MD: Johns Hopkins University, Center for Social Organization of Schools, 1978.

Karweit, N. *Time-On-Task: Issues of timing, sampling and definiton.* Unpublished manuscript, Johns Hopkins University, Center for Social Organization of Schools, 1980.

Karweit, N. L., & Slavin, R. E. *Measuring time on-task: Issues of timing, sampling, and definition.* Paper presented at the annual meeting of the American Educational Research Association, Boston, April 1980.

Kirschner Associates, Inc. *Instructional Dimensions Study: Final Report.* Washington, DC, 1977.

Larrivee, B., & Cook, L. *The effect of institutional variables on teacher's attitude toward mainstreaming.* Paper presented at the annual meeting of the American Educational Research Association, Toronto, March 1978.

Larrivee, B., & Vacca, J. M. *A comparison of academic learning time (ALT) for mainstreamed, low, average, and high ability students.* Paper presented at the annual meeting of the American Educational Research Association, Boston, April 1980.

Leinhardt, G. _An investigation of the effects of time._ Pittsburgh, PA: University of Pittsburgh, Learning Research and Development Center, 1974. (Mimeo)

Leinhardt, G. Outcomes of a strategy for program evaluation. _Science Education_, 1977, _61_(1), 1-17.

Leinhardt, G. Opportunity to learn. In _Perspectives on the Instructional Dimensions Study_, Washington, DC: HEW, November, 1978.

Leinhardt, G. Modeling and measuring educational treatment in evaluation. _Review of Educational Research_, 1980, _50_(3), 393-420.

Leinhardt, G., & Seewald, A. M. Student-level observation of beginning reading. _Journal of Educational Psychology_, 1981. (in press)

Leinhardt, G., Zigmond, N., & Cooley, W. W. Reading instruction and its effects. _American Educational Research Journal_, 1981. (in press)

Lomax, R. G. _A generalizability study of the classroom observations of learning-disabled students_. Paper presented at the annual meeting of the American Educational Research Association, Boston, April 1980.

Lomax, R. G., & Cooley, W. W. _The student achievement-instructional time relationship_. Paper presented at the annual meeting of the American Educational Research Association, San Francisco, April 1979.

MacDonald, F. J., & Elias, P. _Beginning Teacher Evaluation Study_. (Phase II. Technical Summary, Final Report). Princeton, NJ: Educational Testing Service, 1976.

Madden, R., Gardner, E. F., Rudman, H. C., Karlsen, B., & Merwin, J. C. Stanford Achievement Test. New York: Harcourt Brace Jovanovich, Inc., 1972.

Marliave, R. Beyond engaged time: Approximations of task appropriateness in terms of ongoing student learning behaviors. Paper presented at the annual meeting of the American Educational Research Association, Boston, April 1980.

Marliave, R., Fisher, C., Filby, N., & Dishaw, M. The development of instrumentation for a field study of teaching. San Francisco, CA: Far West Laboratory for Educational Research and Development, 1977.

Revicki, D., & Rubin, R. Models for measuring program implementation: A review and critique. Paper presented at the annual meeting of the American Educational Research Association, Boston, April 1980.

Rosenshine, B. Evaluation of classroom instruction. Review of Educational Research, 1970, 40(2), 279-299.

Rosenshine, B. Curriculum and other contextual variables. Paper presented at the annual meeting of the American Educational Research Association, Chicago, April 1974.

Rosenshine, B. Classroom instruction. In N. L. Gage (Ed.), The psychology of teaching methods. Seventy-Seventh Yearbook of the National Society for the Study of Education. Chicago: University of Chicago Press, 1976.

Rosenshine, B. V. Academic engaged time, content covered, and direct instruction. Paper presented at the annual meeting of the American Educational Research Association, Toronto, March 1978.

Rosenshine, B. V. Instructional principles in direct instruction. Paper presented at the annual meeting of the American Educational Research Association, Toronto, March 1978.

Samuels, S. J., & Turnure, J. E. Attention and reading achievement in first-grade boys and girls. Journal of Educational Psychology, 1974, 66(1), 29-32.

Smyth, W. J. Pupil engaged learning time: Concepts, Findings and Implications. The Australian Journal of Education, 1980, 24(3), 225-245.

Smyth, W. J. Research on classroom management: Studies of pupil engaged learning time as a special but instructive case. Journal of Education for Teaching, 1981, 7(2), 127-148.

Smyth, W. J. Research on pupil engaged learning time: Past, present and future! Paper presented at the Invitational Conference on Research on Instructional Time, Evanston, May 1981.

Stallings, J. A. Allocated academic learning time revisited, or beyond time on task. Paper presented at the annual meeting of the American Educational Research Association, Boston, April 1980.

Webb, N. L., & Nerenz, A. G. The relationship of means of instruction and instructional time to reading skills outcome. Paper presented at the annual meeting of the American Educational Research Association, Boston, April 1980.

Welch, W. W., & Bridgham, R. G. Physics achievement gains as a function of teaching duration. School Science and Mathematics, 1968, 68, 449-455.

Wiley, D. E. Another hour, another day: Quantity of schooling, a potent path for policy. Studies of Educative Processes, Report No. 3, University of Chicago, 1973.

Wiley, D. E. Student achievement measures as Title I eligibility criteria: An overview of the policy study. Paper presented at the annual meeting of the American Educational Research Association, New York, April 1977.

Wiley, D. E., & Harnischfeger, A. Explosion of a myth: Quantity of schooling and exposure to instruction, major educational vehicles. Educational Researcher, 1974, 3(4), 7-12.

Wiley, D. E., & Harnischfeger, A. Distinct pupils, distinctive schooling: Individual differences in exposure to instructional activities. Paper presented at the annual meeting of the American Educational Research Association, Washington, March-April 1975.

Table 1

Science Time Study

Combined 4th and 5th Grades (N = 86)

Correlation Matrix

	Combined Input	Science Time	Math Time	Reading Time	Combined Outcome	With Input Partialled Out
Input						
Science Time	-.05					.37
Math Time	.36	.07				.07
Reading Time	.09	.10	.62			-.03
Combined Outcome	.90	.11	.35	.06		
Mean	17.99	1679.83	5405.97	1768.27	16.75	
S.D.	4.71	536.50	1534.00	801.53	5.07	

Outcome = .89*PRE + .16*ST + .07MT − .08RT

Adjusted Mult R^2 = .83*

*$p \leq .05$

Figure 1

Model of Classroom Processes

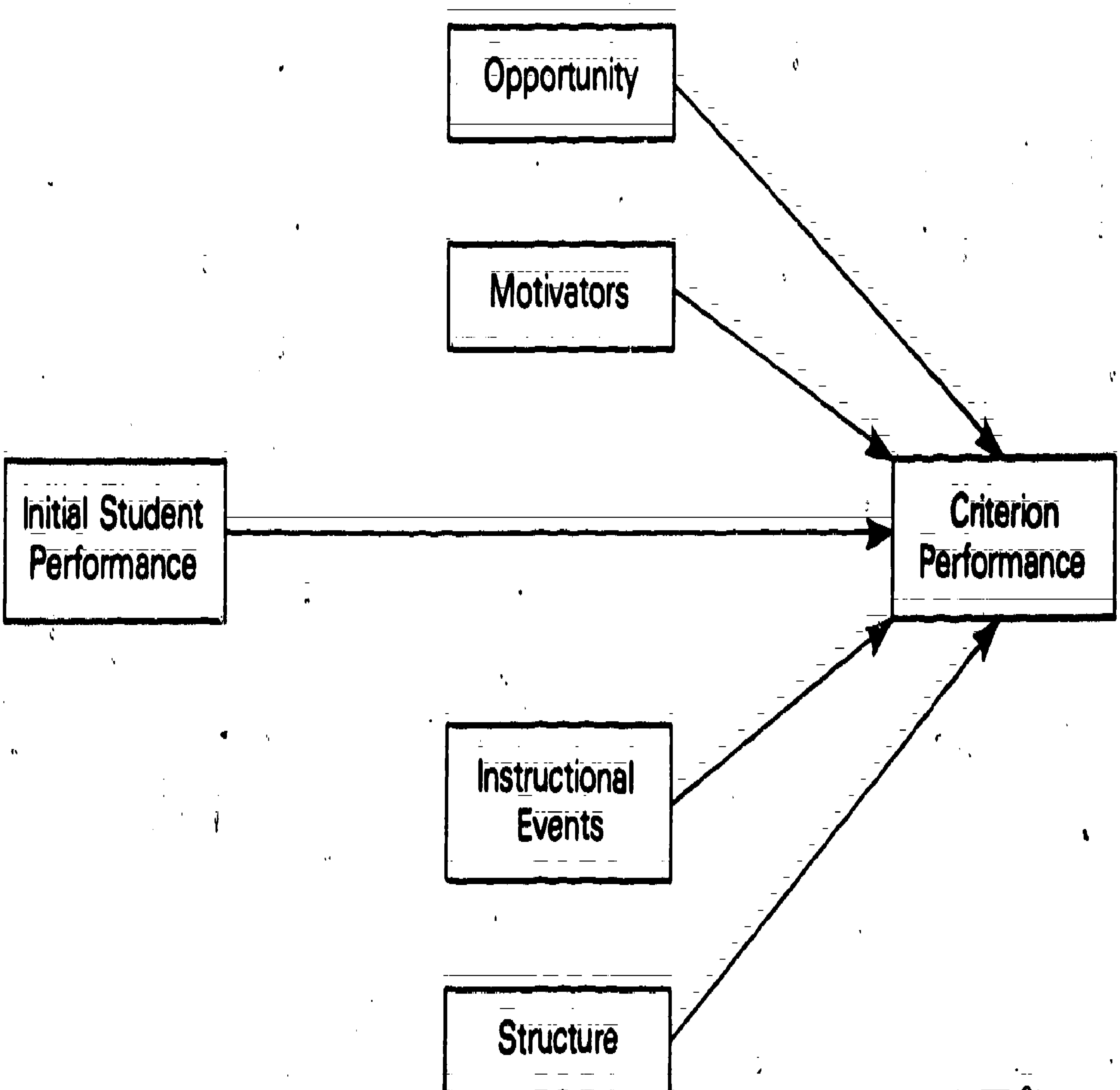
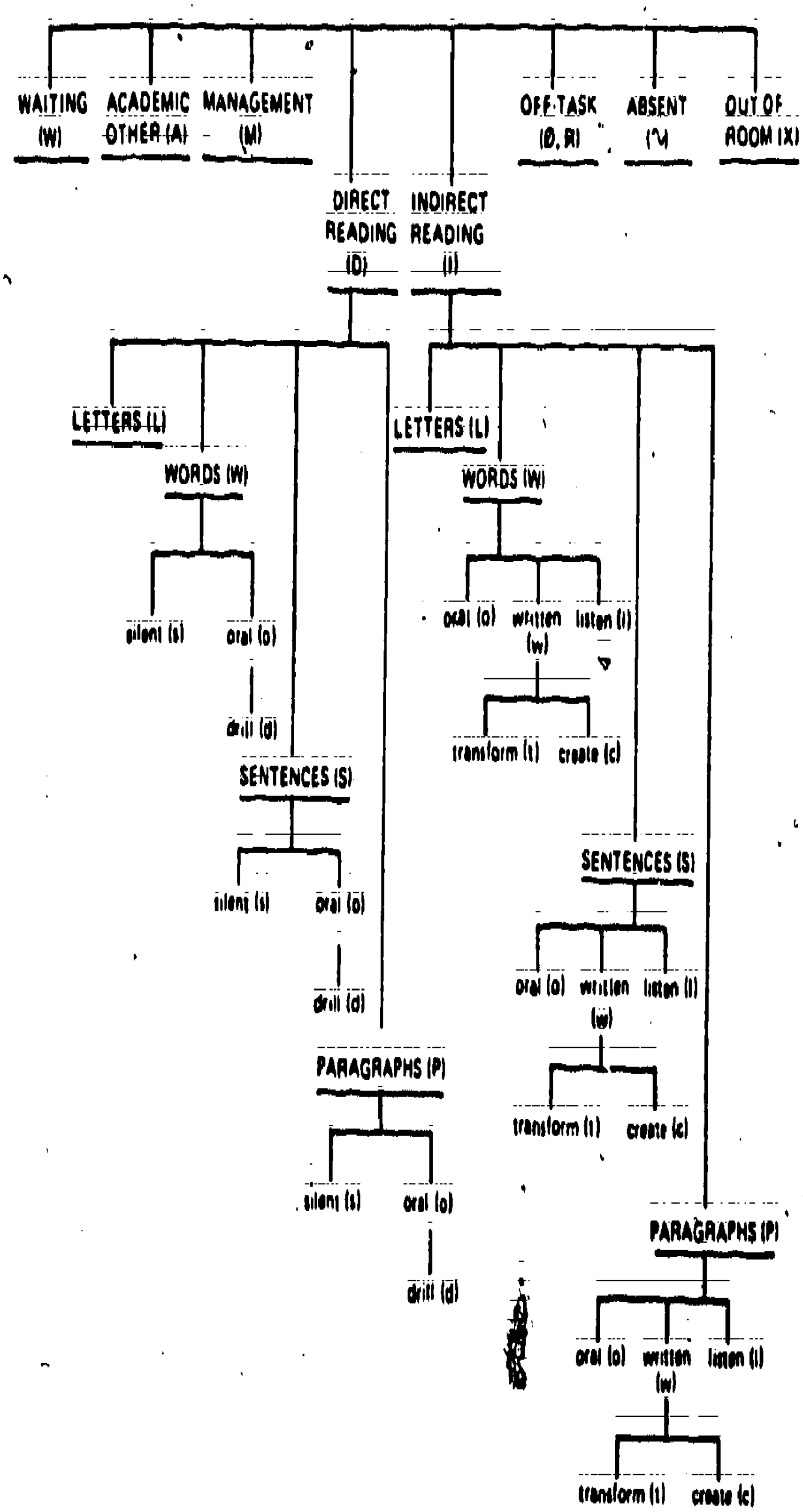


Figure 2

Student Observation of Beginning Reading



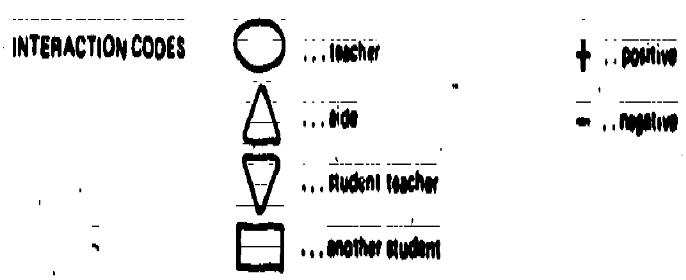
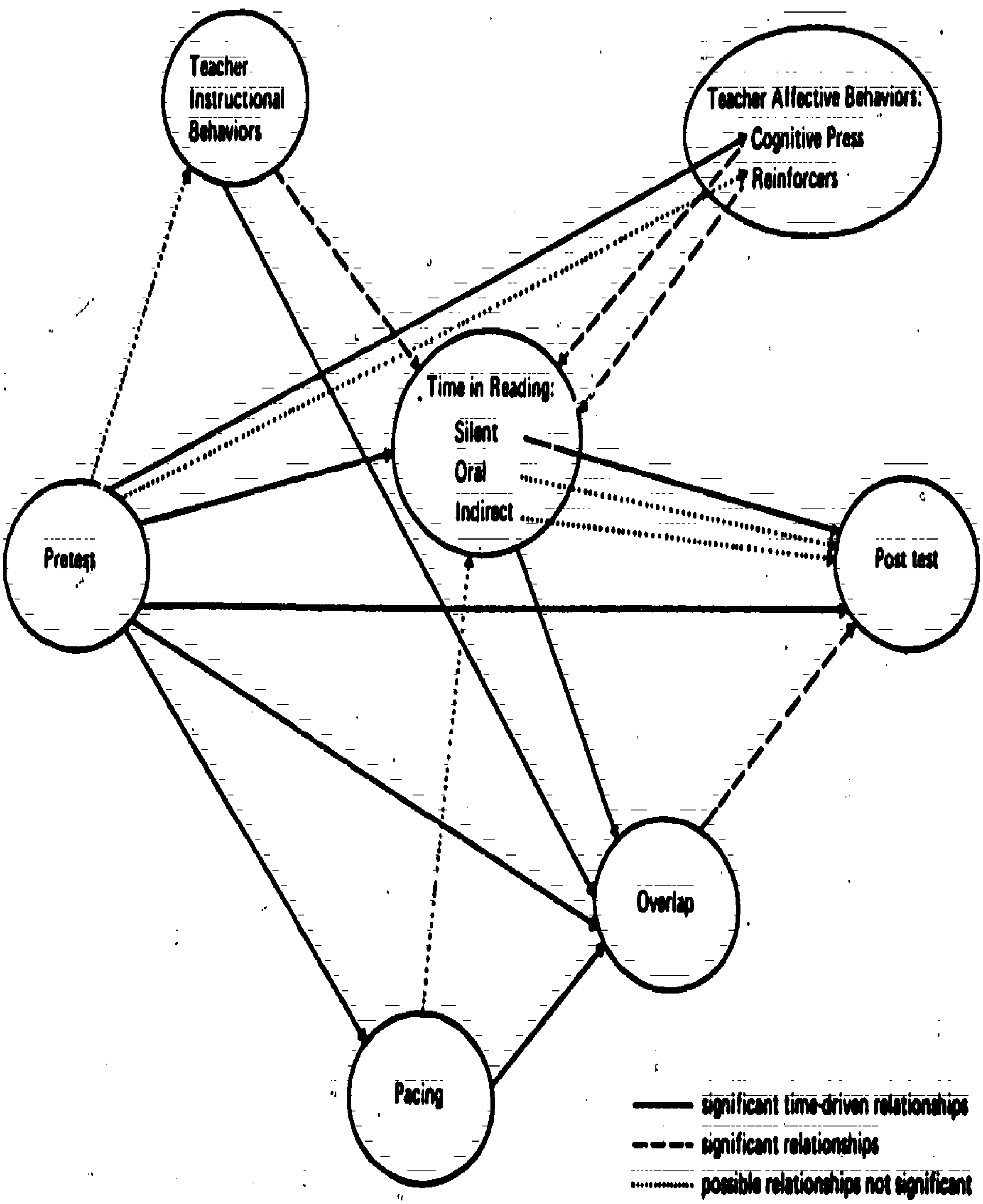


Figure 3

A Model of Reading Instruction



Posttest = 176.6 + 6.2*PT + 1.0*SR + .5OR – .09IR + .4*OV   Adjusted R^2 = .72*

Time in Reading = 13.7 + .7*PT + 1.0*TI + 5.0*CP + .04*R + 2.5P   Adjusted R^2 = .59*

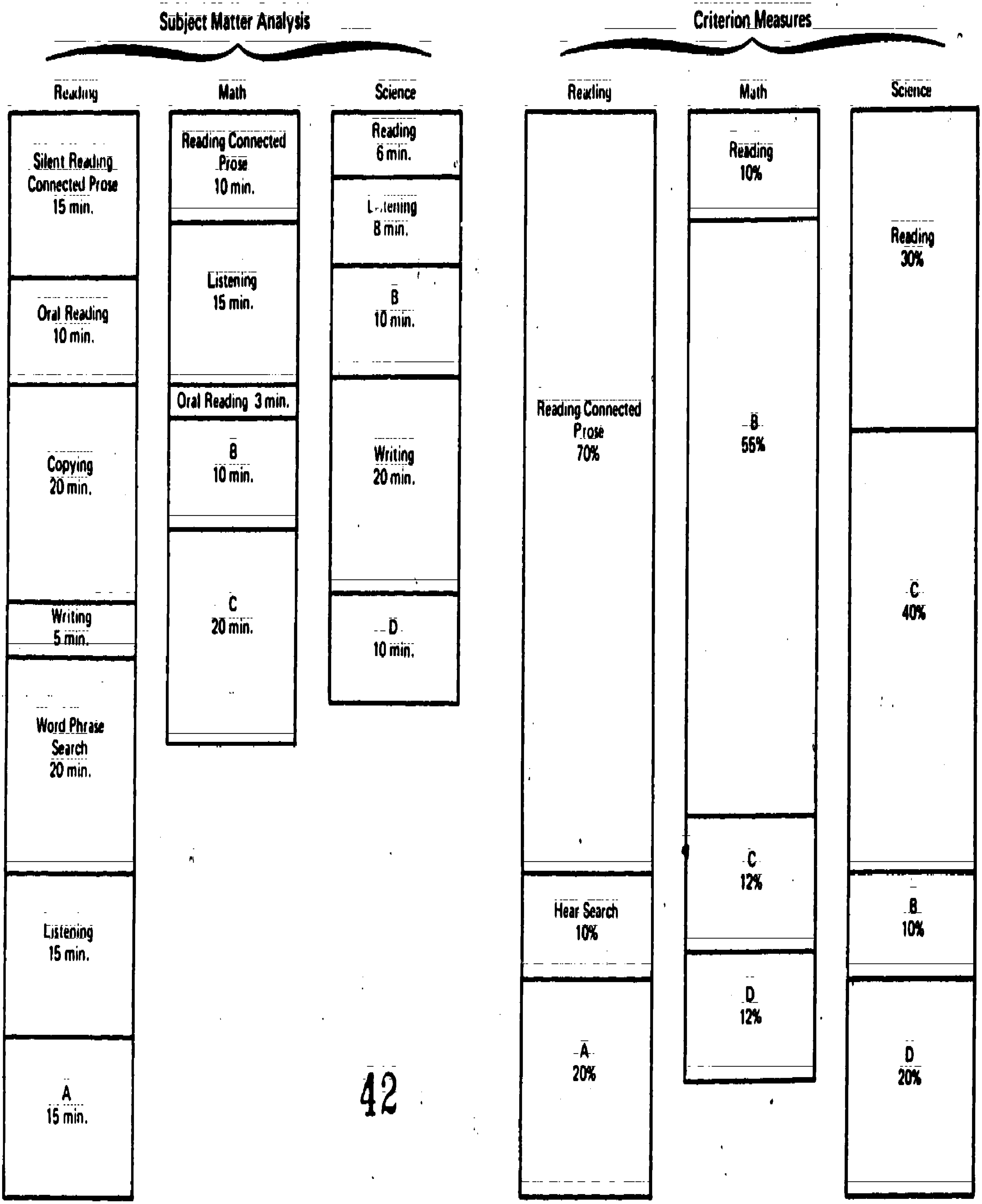


Figure 4

Hypothetical Analysis of Student Time in Subject Matter and Criterion Demands