The monograph describes an approach to special education research that addresses both the promise of immediate payoff for decision makers inherent in program evaluation/validation, and the need to identify effective intervention variables of model programs that could be incorporated into other service settings. The approach, termed "program component research," is illustrated in a successful Child Service Demonstration Center for Children with Learning Disabilities. The application of program component research demonstrates that experimental research can be integrated within existing programs in such a way to benefit both research and service. (CL)
EXPERIMENTAL ANALYSIS OF PROGRAM COMPONENTS
AN APPROACH TO RESEARCH IN CSDC'S

Stanley L. Deno, Berttram Chiang, Gerald Tindal, and Marie Blackburn

University of Minnesota

Research Report No. 12

Institute for Research on Learning Disabilities

STANLEY L. DENO, BERTTRAM CHIANG, GERALD TINDAL, AND MARIE BLACKBURN

INSTITUTE FOR RESEARCH ON LEARNING DISABILITIES

Permission to reproduce this material has been granted by
James Vosselbyke

to the Educational Resources Information Center (ERIC)
The Institute for Research on Learning Disabilities is supported by a contract (300-77-0491) with the Bureau of Education for the Handicapped, Department of Health, Education, and Welfare, U.S. Office of Education, through Title VI-C of Public Law 91-230. Institute investigators are conducting research on the assessment/decision-making/intervention process as it relates to learning disabled children. Research activities are organized into eight major areas:

I. Adequacy of Norm-Referenced Data for Prediction of Success

II. Computer Simulation Research on the Assessment/Decision-making/Intervention Process

III. Comparative Research on Children Labeled LD and Children Falling Academically but not Labeled LD

IV. Surveys on In-the-Field Assessment/Decision Making, and Intervention

V. Pathological Research on Placement Team Decision Making

VI. Bias Following Assessment

VII. Reliability and Validity of Formative Evaluation Procedures

VIII. Data-Utilization Systems in Instructional Programming

Additional information on these research areas may be obtained by writing to the Editor at the Institute.

The research reported herein was conducted under government sponsorship. Contractors are encouraged to express freely their professional judgment in the conduct of the project. Points of view or opinions stated do not, therefore, necessarily represent official position of the Bureau of Education for the Handicapped.
Research Report No. 12

EXPERIMENTAL ANALYSIS OF PROGRAM COMPONENTS:
AN APPROACH TO RESEARCH IN CSDC'S

Stanley L. Deno  Berttram Chiang  Gerald Tindal
University of Minnesota

Marie Blackburn
Minneapolis Public Schools

Institute for Research on Learning Disabilities
University of Minnesota

August, 1979
Experimental Analysis of Program Components
An Approach to Research in CSDC's:

The primary purpose of the program to develop Child Service Demonstration Centers for Children with Learning Disabilities (CSDCs) (Title VI C) was to provide a system for disseminating programs which were demonstrably effective. From the outset, it could have been argued that few, if any, instructional programs existed for which effectiveness had been empirically demonstrated, and that research rather than development was what should have been funded. An alternative to this argument was that an empirical base for dissemination of a program could be developed by requiring CSDCs to undergo a "validation" procedure (Tallmadge, 1977). Validation involved collection of evaluation data in which the achievement of students served through the CSDC was contrasted with the achievement of students in other educational programs. If the comparison was carefully done and CSDC students' performance exceeded that of comparison students, then program effectiveness was said to be empirically "validated" and could legitimately be disseminated.

While methodological grounds could certainly be found for questioning the internal and external validity of individual validation attempts, the potential for identifying effective programs through validation certainly exists. Special education programs in general, and model programs in particular, should be encouraged to continually obtain the best
possible data on their effectiveness, since their very existence is predicated on the assumption that the additional cost of such programs is offset by increased benefits to the students. Program validation, however, is in the tradition of efficacy research, and adequate efficacy research is difficult, if not impossible, to do (Cegelka & Tyler, 1970). Even when done well, efficacy research is typically based on what Gelfand and Hartmann (1975) call "blockbuster interventions." Blockbuster interventions are those in which a complex bundle of treatment variables is constructed by the program designers to yield the highest probability of impacting the problem.

CSDC programs clearly can be classified as blockbuster interventions. Commendably, administrators and teachers do all they can to ensure the effectiveness of their program. The problem for potential adopters of a CSDC program, however, is to know which of all the components of the program are really essential to its success. Is it the curriculum, the staff competence, the staff-student ratio, the parent involvement, the instructional methodology, or some interaction among all or a subset of those components which is critical to program success? The question is not simply academic, since variations in program components have economic effects.

An alternative or complementary approach to evaluating blockbuster interventions which can simultaneously yield the necessary information for decision makers and increase our knowledge concerning the effective variables in treatment is research on program components.

Program component research, in contrast to typical program efficacy research, is designed to yield data on the effects of varying specific components within a program rather than on the effects between programs.
in general. Program component research is much the same as research on instruction in which variations in methodology are empirically analyzed. It differs in two important respects, however. First, the primary goal of program component research is to construct a more efficient and effective service program. Second, program component research is typically conducted within the context and constraints of existing educational programs. In these two respects, program component research is similar to research on instructional theory as recently outlined by Snow (1977), and it is clear that research of this type places certain constraints on research design.

Since the primary purpose of program component research is to construct a better program, and the research is conducted within an existing program, the selection of treatment contrasts is guided by what variables can be manipulated within that program. If a program includes instruction in decoding words, and teachers are free to vary their instruction in word decoding, then research can be organized to determine the effects of such variations. Were decoding instruction either not a program component or a component which could not be manipulated, then research would not be designed to test the contribution of instruction in decoding to program success.

While limiting research questions to those relevant within a particular instructional program may seem to confine the research investigators, the truth is that most of the variables of interest to researchers are present in any program. What tends to be limiting is how the variables are operationalized by both the staff and the existing curriculum. At the same time, the constraints of program component research are offset by the direct applicability of the outcomes.
staff can use the results to modify practice without needing to question its relevance to their program. Then too, if the student population and setting are clearly described, other program administrators interested in adopting a program will have data on the relative value of individual program components.

To be successful, program component research requires program staff who are open and inquisitive about what they are doing. Clearly, such research implies that all of what has been included in a program may not be necessary or helpful. Too often special educators are like Campbell's (1969) "trapped administrators," required to act as if they know what they are doing is correct. Program component research can free program administrators from that trap by operating on the premise that special education is "experimental education" (Burelllo, Tracy, & Schultz, 1973), to be used as "developmental capital" (Deno, 1970) for improving the capacity of educators to serve individual needs.

A good illustration of program component research in special education is reported by Jenkins and Mayhall (1976). At the time of the research the authors were involved in developing the Remediation Specialist Program through the Southwest Special Education Regional Resource Center. In an effort to develop an empirical basis for Remediation Specialist functions, Jenkins and his colleagues experimentally determined the effects of such program variables as peer and cross-age tutoring, tutor-pupil ratio, site and supervisor of tutoring, and frequency of tutoring. The results of their research provided them with the data necessary for selecting effective program components, and also provided the field of special education, in general, with guidelines for effective resource programming (Sindelar & Deno, 1977).
Program Component Research in a CSDC

In 1975, the Bureau of Education for the Handicapped awarded a Title VII grant to the Minneapolis Public Schools to develop a Child Service Demonstration Center for Children with Learning Disabilities. Subsequently this project, eventually known as SIMS (Systematic Instructional Management Strategies), became highly successful in disseminating its program both regionally and nationally, and underwent validation procedures which provided evidence that the program, in general, was effective. In its final year of funding, the project director sought to improve program dissemination by developing a research component which might yield data useful both to the project and to the field of learning disabilities. The result was a one-year effort to conduct a series of studies on the components of the SIMS program.

Selection of Program Components for Research

Since a program like SIMS consists of a large number of interacting components, selecting those components on which to do the research in the allotted time was a difficult task. As indicated earlier, a balance had to be achieved between what program variables might be most important to evaluate, and what program variables could be evaluated within the constraints created by an existing service program. The eventual selection of program components was determined by selecting those which:

(a) were unique to SIMS, (b) had available on them little research evidence, (c) could be varied once the school year had begun, (d) were adaptable to a research design enabling sufficient experimental control, and (e) could be incorporated in a research design that would be minimally intrusive with respect to routine staff activities. Using these selection criteria, the following program components were included:
1) The daily data collection procedures used to monitor
students' progress on the SIMS reading and spelling skill/concept sequence.

2) The data utilization techniques based on the daily data
collection system.

3) The instructional techniques used by the teachers in teaching
the SIMS reading and spelling skills/concepts.

Specific Research Questions. Just as a service program is a set
of components, the SIMS components identified for research were, in
fact, a complex of events for which a number of research questions could
be generated. The research had to be designed to answer a limited set
of specific questions with respect to each component. Those questions
were:

1. With respect to the daily data collection procedures, does
graphing student performance data obtained through daily
timed testing increase rate of mastery beyond what occurs
when simply maintaining the data on recording sheets?

2. With respect to data utilization, does using a decision
rule based on an aim line (Liberty, 1975) to make tokens
contingent increase mastery beyond simply setting aim lines
without decision rules and tokens?

3. With respect to instructional techniques, does direct
instruction in phonetic rules increase mastery beyond drill
and practice on phonetically regular words without instruction
in rules?
Method

All of the research was conducted with the students enrolled in the Systematic Instructional Management Strategies (SIMS) Centers at the elementary and secondary levels. The SIMS Centers provide instruction in basic language arts and math skills to those students, grades 1-9, who have not responded to programs offered in resource rooms at their home schools.

Of the 32 students (18 elementary and 14 secondary) who participated in the research, two were females. Most of the students had been in the SIMS program for a period of one to two years. A few students did not participate in the research either because they had not entered the reading and spelling sequence, or they were progressing too slowly to fulfill research design requirements.

Experimental Design. A within-subject design was used for all phases of the research. The design was created by having each student experience each treatment on a different skill/concept category in the reading and spelling sequence. A student who was on Category #6 in the sequence might experience Treatment A on that category, and Treatment B on Category #7. A second student who was in Category #12 might receive Treatment B on that category and Treatment A on Category #13. The design was chosen because it introduced experimental controls least intrusive to routine instruction. In addition, inter-class and inter-subject variability was reduced to a minimum since all subjects were assigned to all treatment conditions. Treatment conditions were assigned to random categories to control for sequence effects and the different levels of category difficulty.
The experimental research was organized in three phases. The phases were designed to contrast: (a) graphing daily performance versus no graphing of daily performance, (b) using aimlines and decision rules with a token economy versus aimlines and decision rules; and (c) drill and practice versus conceptual teaching. An additional comparison was made across phases between graphing with aimlines and graphing without aimlines.

Materials. All of the research was conducted on reading. The reading materials consisted of the isolated word lists developed by the SIMS program. The words on these lists follow exclusively the word patterns for the 53 designated skill/concept categories.

The other materials used were recording sheets, stopwatches, and semi-logarithmic graph paper for graphing student performance. During the second and third phases, the points were given as token reinforcers could be exchanged primarily for free-time choice activities, and other items such as soft drinks, posters, etc. In the third phase of the study, instructional activities typically included the use of word and syllable cards.

Procedures. For each experimental phase, a treatment script was developed to outline all the specific procedures to be followed and to ensure consistency across all subjects. Within each phase, the order of treatment conditions was randomly counterbalanced. The standard daily timed-test procedure was established using isolated word lists associated with each category. A student was given a one-minute practice test which was followed by a one-minute timed test. The procedure was essentially the same as used routinely by all the teachers. The number of words read correctly and the number of words incorrect
from the second timed test were recorded.

Students ordinarily practiced initially on decoding words from a given category without timed testing until they could identify words with 90% accuracy. A skill/concept category was used for the research if a student had not done timed testing previously on that category. Students were tested each day throughout the research. Treatments for a given category were continued until a student attained a mastery criterion on 40 words per minute correct on two of three days, or until a period of 15 school days elapsed, whichever occurred first. During the first two phases of the research, as was routine, the daily timed tests of performance on the isolated word lists were administered exclusively by the instructional aides. The aides continued the testing during Phase I, but the experimental treatments (variations in instruction) were accomplished by the teachers.

Phase I (Graphing vs. no graphing): During this first comparison, 30 students (18 at Armatage and 12 at Franklin) had their daily timed test score graphed and not graphed. Under the graphing condition, the results of 1-minute timed tests were recorded on a raw data sheet and graphed on six-cycle semi-logarithmic graph paper. Under the no graphing condition, the results were recorded only on raw data sheets. Under both conditions, the aides were instructed to give praise for improved performance.

Phase II (Decision rule and tokens vs. decision rule and praise): For the second experimental comparison, the graphing procedures described above were maintained, and a data-decision rule system using aim lines was implemented to determine when points for improvement
were to be delivered. For each subject, a baseline was established by using the median of three times on the first day. This data point served as a starting point for drawing an aim line on the graph. A predetermined 30% increase per week over the baseline was used as the slope for the aim lines. When students were under the data-decision-rule condition, data utilization was accomplished each day by giving one point for reaching the aim line and one extra point for every word per minute achieved above the aim line. An attempt was made to standardize point values across all classrooms. A decision rule was used for shifting the aim line whenever the subject's performance fell below the line three days in a row. In this instance, the median level of these three days was used to reset a parallel aim line (i.e., the slope of a 30% increase remained the same). Each time the aim line was shifted, the number of points provided for reaching aim line was increased by one. To prevent the subjects from performing poorly in order to increase point earnings, a 10-point bonus was provided for reaching the aim line five days in a row. When students performed under data utilization with social consequences, the establishment and adjustment of aim lines was the same as previously described. However, only visual and verbal feedback on performance and praise were provided by the aides.

Phase III (Concept teaching vs. drill and practice): Two instructional strategies were contrasted during this phase — drill and practice and what the SIMS staff called "concept teaching." All previously described procedures involving graphing, decision rules, and points were maintained during this phase. Each day the teacher provided a ten-minute period of instruction, directed toward the category in which the timed
tests were given. The instruction was given prior to testing. The drill and practice condition consisted of repetitions of phonetically regular words from the relevant category. The words were presented in isolation and in context with error correction, but no direct instruction was provided in the phonetic (conceptual) rules which underlie category organization. Practice with flash cards, word lists, or T-scope were the primary modes of presentation. Concept teaching, on the other hand, focused on direct instruction in the phonetic rules underlying the words. Teachers provided direct instruction in word analysis and synthesis through verbal explanation and sorting word parts and whole words on the basis of the common phonetic rules for decoding words within the category.

Results

For all three phases, the results were analyzed using the number of days taken to reach the mastery criterion. Table 1 contains the results in terms of the mean number of days taken to criterion. The results were statistically analyzed for the elementary and secondary students separately using both correlated t tests and the Wilcoxon Matched Pair Signed Ranks Test. The outcomes for both analyses also are presented in Table 1.

---

Insert Table 1 about here

---

Phase I (Graphing vs. no graphing): The analyses with respect to graphing produced mixed results. A statistically significant difference was obtained using the Wilcoxon Test with the elementary student data, but not the junior high data. This result means that, among the
elementary students, a significantly higher proportion attained mastery more rapidly when their daily performance was graphed than when it was not graphed. When a correlated t test was conducted on the mean number of days required to achieve mastery under each of the treatment conditions, no difference was obtained at either the elementary or secondary level.

Phase II (Decision rules and points vs. decision rules and social consequences): When student performance under the two decision rule conditions was examined, significant effects were obtained for both mean days to criterion (correlated t) and the proportion of students performing better (Wilcoxon T) under the data-decision rule with points condition. Again, however, the effect was obtained at elementary level but not at the junior high level.

Phase III (Concept teaching vs. drill and practice): By the end of the school year only four students completed this phase at the elementary level, so the majority of the data for this analysis was from the junior high level. The Wilcoxon analysis revealed that a higher proportion of students performed better under concept teaching at the secondary site. Results from the correlated t analysis revealed no difference in mean number of days to criterion under the two teaching conditions. When the data from elementary and secondary sites were combined, the same results were obtained.

Cross-Phase Comparison: An attempt was made to evaluate the effect of using graphs with an aim line in contrast to graphing without an aim line. To make this comparison, data from the graphing condition of Phase I were compared with data from the decision rule and praise condition of Phase II. The results of this comparison revealed that a
higher proportion of students at the elementary site achieved criterion faster when daily performance was simply graphed without aim lines. No difference was obtained on the mean days to criterion under the two conditions.

Discussion

The experimental analysis of SIMS' program components yielded several useful outcomes. One of the most noticeable findings is the different effects of the treatment conditions between the two sites. A difference in student performance was obtained for each paired comparison of program components. Three of the effects were obtained at the elementary site and one at the junior high site. The one significant effect at the junior high site occurred in the comparison between drill and practice and concept teaching, while the small number of students completing that phase at the elementary site did not permit an adequate test of that component for the elementary students. The difference between the two sites could be related to several factors. One of the most obvious, of course, involves differences in age. Not only are students in junior high older than those in elementary school, but they also have experienced a longer history of reading difficulties and intervention attempts. The isolated word lists which were used for both testing and practice often consisted of very simple words categorized into similar patterns. The older students may have had more difficulty persisting on what they recognized as a low level reading task. Some students at the junior high site showed a less positive attitude toward reading the word lists and took part only reluctantly. This is, of course, a program component which could be experimentally analyzed.
In addition to the difference in outcomes for the elementary and junior high programs, several important results were obtained. First, how the daily measurement data are recorded and displayed for students at the elementary level appeared to be important. Students generally attained mastery faster under conditions where they and the aides could see daily performance on a graph. A related finding was that, for the elementary students, adding aim lines with decision rules for resetting the aim lines and praise for attainment appears not to increase the rate of achievement. That conclusion, however, is based on a comparison of performance between categories to which treatments were not randomly assigned; therefore, tentativeness seems warranted. The results, in general, provide support for including daily graphing of student performance as a component of an intensive instructional program for elementary-aged learning disabled students.

With respect to the second research question, the reasonable conclusion parallels the first outcome. Clear support was obtained for the effectiveness of using a data-decision rule system to deliver tokens for achievement. Again, however, the effect held with elementary aged students rather than the junior high students. The failure to obtain an effect for data-decision rules with tokens at the junior high level is most likely due to either the particular data-decision rule used or the token exchange system and its backup consequences. Since a plethora of data exists on the effectiveness of token systems, it seems unwarranted to conclude that the results provide evidence that the token economy component should not be included in the junior high program. Instead, the results provide a basis for suggesting that more attention be given to identifying the appropriate decision rules for the secondary
students, and to determining how the token economy can be made more effective. The answer to the second research question would seem to be that teachers incorporating SIMS into their programs for elementary aged children might well include data-decision rules for delivering tokens to increase their effectiveness.

Finally, with respect to the relative value of the different teaching approaches, the evidence from this analysis provides support for direct instruction in the teaching of conceptual rules. It is interesting that this conclusion is based on the performance of the secondary students, since propositional logic is something of which they are developmentally capable. Unfortunately, the implication that an age by treatment interaction exists is not supportable given the small number of elementary students completing these treatments. Beyond our program component analysis, the value of practice in applying phonetic rules to decoding is, nevertheless, supported by the results of the research. This finding also is intriguing because the daily timed testing procedure used tended to penalize "sounding out" or verbal self-instructions as an aid to performance. Apparently, the effectiveness of conceptual teaching accrues during the teaching time rather than the testing time. The results support the recommendation that direct teaching of concepts remain a part of the SIMS program.

While the results of the experimental research support graphing of daily performance, using decision rules for token delivery, and direct teaching of concepts as valuable components of the SIMS program, several caveats should be issued. First, all the research was conducted on components that require a small fraction of the time devoted to
reading instruction. The proportion of total variation in reading achievement accounted for by these components cannot be ascertained through the research that was conducted. A second concern is that all the conclusions rest entirely on the mastery criterion based on saying isolated words from category word lists. Since the criterion validity of this measure is not adequately established, the importance of the results for other dimensions of reading performance such as comprehension and silent reading is unknown.

**Summary**

In this paper, we have attempted to describe an approach to research in special education that addresses both the promise of immediate payoff for decision makers inherent in program evaluation/validation, and the need to identify those effective intervention variables embedded within model programs that might be incorporated into the services delivered by professionals in other settings. We have labeled this approach "program component research," and have provided an illustration based on research in one successful Child Service Demonstration Center for Children with Learning Disabilities. The research results presented in our illustration have provided data specifically useful to the decision makers in that program. At the same time, however, the results could be useful to other program administrators and to researchers interested in the variables which we experimentally manipulated.

What may be more important than the particular research results, however, is the demonstration that experimental research can be integrated within existing service programs in a way that directs and
benefits both research and service. The tension experienced by federal research funding agencies over what priority to give program development as opposed to fundamental research might to some degree be reduced if priority is given to research on program components. To do so will require a sustained commitment of funds to collaborative efforts between research and service institutions.
References


Deno, E. Special education as developmental capital. Exceptional Children, 1970, 37, 229-237.


Footnote

Berttram Chiang is now Assistant Professor in the Special Education Program, College of Education at the University of Wisconsin, Oshkosh. Maxie Blackburn is affiliated with the Institute for Research on Learning Disabilities as a school liaison.
Table 1

Number of Days Taken to Reach the Mastery Criterion.

<table>
<thead>
<tr>
<th>Site</th>
<th>Component</th>
<th>Elementary correlated Wilcoxon</th>
<th>Junior High correlated Wilcoxon</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>N</td>
<td>X</td>
</tr>
<tr>
<td>Phase I</td>
<td>graphing</td>
<td>18</td>
<td>11.33</td>
</tr>
<tr>
<td></td>
<td>no graphing</td>
<td>18</td>
<td>11.94</td>
</tr>
<tr>
<td>Phase II</td>
<td>token consequence/decision rule</td>
<td>14</td>
<td>9.36</td>
</tr>
<tr>
<td></td>
<td>social consequence/decision rule</td>
<td>14</td>
<td>11.00</td>
</tr>
<tr>
<td>Phase III</td>
<td>drill and practice</td>
<td>4</td>
<td>10.25</td>
</tr>
<tr>
<td></td>
<td>concept teaching</td>
<td>4</td>
<td>10.00</td>
</tr>
<tr>
<td>Cross-Phase Comparison</td>
<td>graphing without aimline</td>
<td>11</td>
<td>11.36</td>
</tr>
<tr>
<td></td>
<td>graphing with aimline</td>
<td>11</td>
<td>12.55</td>
</tr>
</tbody>
</table>

*p < .05  **p < .01
PUBLICATIONS

Institute for Research on Learning Disabilities
University of Minnesota

The Institute is not funded for the distribution of its publications. Publications may be obtained for $2. per document, a fee designed to cover printing and postage costs. Only checks and money orders payable to the University of Minnesota can be accepted. All orders must be prepaid.

Requests should be directed to: Editor
Institute for Research on Learning Disabilities
350 Elliott Hall
75 East River Road
University of Minnesota
Minneapolis, Minnesota 55455


*Ysseldyke, J. E., & Thurlow, M. L. Specific investigations to be completed during years two and three (Monograph No. 4). June, 1978.


*As part of its continuation proposal, the Institute was required to prepare these monographs. Because they are part of the proposal, they are not available for general distribution.

**This research report is not being distributed by the Institute. Requests for it should be directed to: The NETWORK, 290 S. Main Street, Andover, Massachusetts 01810.


