

DOCUMENT RESUME

ED 125 775

PS 008 682

AUTHOR Ferb, Thomas; And Others
 TITLE Comparing Curricular Models in Project Follow Through from Kindergarten to Third Grade.
 INSTITUTION Abt Associates, Inc. Cambridge, Mass.
 SPONS AGENCY Office of Education (LHEW), Washington, D.C.
 PUB DATE 23 Apr 76
 CONTRACT 300-75-0134
 NOTE 49p.; Paper presented at the Annual Meeting of the American Educational Research Association (San Francisco, California, April 19-23, 1976)

EDRS PRICE MF-\$0.83 HC-\$2.06 Plus Postage.

DESCRIPTORS Academic Achievement; Affective Tests; Comparative Analysis; *Curriculum Research; Discovery Learning; *Early Childhood Education; Open Education; Parent Attitudes; *Preschool Programs; Problem Solving; Programed Instruction; *Program Evaluation; *Symposia; Teacher Attitudes

IDENTIFIERS *Project Follow Through

ABSTRACT

This symposium compares the third grade outcome profiles of Follow Through curricular models classified on the basis of their theoretical underpinnings as (1) behaviorally programed, (2) cognitive-discovery, or (3) open classroom. Outcome profiles are presented for academic tests, conceptual problem-solving, and selected affective measures, and curricular differences in parent and teacher attitudes are also reported. In addition, the extent to which similar models differ, the significance of site-to-site variations, and the analytic issues involved in comparing the various models are discussed. (JMB)

 * Documents acquired by ERIC include many informal unpublished *
 * materials not available from other sources. ERIC makes every effort *
 * to obtain the best copy available. Nevertheless, items of marginal *
 * reproducibility are often encountered and this affects the quality *
 * of the microfiche and hardcopy reproductions ERIC makes available *
 * via the ERIC Document Reproduction Service (EDRS). EDRS is not *
 * responsible for the quality of the original document. Reproductions *
 * supplied by EDRS are the best that can be made from the original. *

U.S. DEPARTMENT OF HEALTH
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRO-
DUCED EXACTLY AS RECEIVED FROM
THE PERSON OR ORGANIZATION ORIGIN-
ATING IT. POINTS OF VIEW OR OPINIONS
STATED DO NOT NECESSARILY REPRE-
SENT OFFICIAL NATIONAL INSTITUTE OF
EDUCATION POSITION OR POLICY

ED 125775

COMPARING CURRICULAR MODELS IN PROJECT FOLLOW THROUGH
FROM KINDERGARTEN TO THIRD GRADE

by

Thomas Ferb
John C. Larson, Ph.D.
Linda B. Stebbins, Ph.D.
Nancy Ames

all of

Abt Associates, Inc.
55 Wheeler Street
Cambridge, Mass. 02139

Presented at the Annual Meeting of the American Educational
Research Association at San Francisco, 23 April 1976

PS 008682

These papers are based on research performed by Abt Associates Inc. under
Contract No. 300-75-0134 to the United States Office of Education. This
research is the product of many hands; neither Abt Associates staff nor
the Office of Education is responsible for any errors or omissions; the
authors accept full responsibility for those. The statements in these
papers do not necessarily represent any official position of the Office
of Education.

DESCRIPTION OF THE STUDY

Thomas Ferb

Abt Associates Inc.
55 Wheeler Street
Cambridge, Mass. 02138

In the presentations which follow, three curricular approaches, one Cognitive, one Behavioral, and one Open are presented. Each of the approaches is represented by two Follow Through programs; however, the particular Follow Through programs and the sites involved are not identified by name. The letters A through F are used for the identification of programs and numbers for sites. The programs are not identified because the focus of this symposium is on the individual curriculum and its particular impact. Furthermore, the results presented are based on the Cohort II stream of Follow Through data which is not the primary focus for impact evaluation. The Cohort III stream which is currently being analyzed is the main focus for the evaluation purposes and will be reported on later this year.

In the following presentations, a particular format is followed. The presentations begin with a definition of the curricular approach, followed by a specification of the sample in terms of the particular children and the sites involved, followed by some data on the reported attitudes and practices of teachers, and finally an explication of the impacts of the programs on children is presented in terms of norm scores obtained at the end of third grade and in terms of gains or progress from the beginning of kindergarten to the end of third grade.

In this presentation I will deal with some definitional and interpretive detail concerning the sample characteristics, the teacher measures utilized, the child outcome measures, the analysis strategy, and a major weakness of the study related to Head Start and preschool. Let us begin with the sample. The children utilized for this study participated either in the Follow Through program or one of the selected comparison groups in the Cohort II stream of the Follow Through evaluation. This group began kindergarten in 1970 and completed third grade in 1974. In general, the Follow Through children are substantially below the national population on many socioeconomic status characteristics. The median income of the national

population in 1970 was \$9500, while family median income in Follow Through was only \$4400. This large income difference between the Follow Through children and the national population suggests that the Follow Through program was successful in recruiting a poverty level sample of pupils. Other characteristics of the sample substantiate this general disadvantage. Eighty-one percent of the FT children participated in one or another preschool, whereas only nine percent of kindergarten children in the national population in 1970 participated in preschool. The majority of children in Follow Through (75%) are Black whereas only 11% of the national population is Black. Family size in Follow Through averages six whereas the national population in 1970 averaged only 3.5. Thus the Follow Through group appears to be substantially different from the community at large.

The Follow Through and comparison groups are not nearly so mismatched. The characteristics of the overall Follow Through and comparison group are quite similar. The median income for the Follow Through group, again, is \$4400, whereas the median income for the comparison group is \$5400. Seventy-five percent of the FT group is Black and 66% of the comparison group is Black. A major difference between the groups does appear on preschool. Eighty-one percent of the Follow Through group graduated from Head Start or other preschool, whereas only 57% of the comparison group are preschool graduates. These differences suggest that the Follow Through group is somewhat disadvantaged relative to the comparison group.

It must be pointed out that the disadvantage of the FT group relative to the comparison group is highly variable across sites. In some locales the Follow Through group and comparison group are very similar whereas in others the comparison group has a substantial socioeconomic advantage. The degree to which the Follow Through and comparison group are comparable has some obvious implications in that different groups present different problems, not only in terms of the social and educational problems with which the programs must deal but also in terms of analysis problems related to estimating the progress that the children have achieved over the four-year period of participation in Follow Through. Furthermore, the likelihood that the comparison children participated in some other compensatory program probably varies among sites. This lack of comparability will be considered below when the analytic strategy is described.

Let us now turn to the teacher measures that are utilized. The teacher data reported here deal exclusively with third grade teacher questionnaires, again from the Cohort II group. We make use of five dimensions or clusters which attempt to characterize the learning environment of the children and may serve as proxy variables for program implementation. The clusters include: 1) teacher-centeredness, that is, the degree to which the classroom is controlled and structured by adults; 2) classroom flexibility, that is, the degree to which the children work independently and exercise freedom of choice; 3) individualization of instruction -- the degree of emphasis on small group instruction; 4) academic vs emotional emphasis, that is, the degree to which teachers emphasize basic skills rather than socio-emotional development, and finally, 5) teacher/satisfaction with the sponsor's program, that is, the degrees of the Follow Through teacher's liking for and assessment of the usefulness of the sponsor's program. These dimensions were developed from the Pearson Correlation Matrix of items from the teacher questionnaire. The correlations were submitted to a cluster analysis using an average linkage between merged groups to identify the dimensions or clusters. The methods of reciprocal averages was then employed to develop item weights for each item response in each of the cluster variables. The internal consistency of the scales was assessed and they were found to have adequate reliability. Finally, the cluster scores were standardized across the Follow Through and comparison group teachers using a mean equal to 50 and a standard deviation equal to 10. Note that these data are being used to get a sense of program implementation within a site. The different curricular approaches, Behavioral, Open, and Cognitive, have very different objectives with regard to the learning environment and hopefully these five dimensions characterize the goals and differences.

Please note that Figure 1 shows a T score for each of these teacher dimensions -- teacher centeredness, classroom flexibility, individualization of instruction, emotional vs academic emphasis, and teacher satisfaction within the program; a value is shown for each of the sites in which a particular program is operating. The sites are differentiated by the key at the top of the Figure. Since the scores have been standardized, differences between the value for a particular site and the overall average can be directly interpreted from the table in terms of T score differences. Our discussion is restricted to differences that are equal to or greater than one-

quarter of a standard deviation, that is, 2.5 T score units. All of the scales are ordered so that higher scores indicate a higher degree of the dimension. Thus, a site with a value on teacher centeredness of 60 may be more structured than a site with an average value of 50. The tentative nature of these data must be emphasized. First of all, the data are based only on third grade teachers. Secondly, within site the number of teachers is sometimes quite low. Furthermore, since the data are based strictly on teacher reports, and not on independent observation their use for either characterizing the learning environment of children or as a proxy variable for the implementation of a particular program may be limited.

Let us now turn to the child outcome measures. A battery of existing standardized tests was chosen and modified for the national evaluation outcome measures. After careful consideration, this battery was selected as a best compromise between the needs for accountability and the difficulty of measuring sponsors' diverse goals and objectives. These measures allow comparisons to be made between the Follow Through and comparison groups on the attainment of basic skills, cognitive and conceptual skills, and affective measures including both self-esteem and attitude toward school and learning. The measures include the Metropolitan Achievement Test with its Spelling, Reading, and Mathematics standard subtests; the Raven's Coloured Progressive Matrices, which measure a child's problem-solving ability in visual and perceptual tasks; and the Coopersmith Self-Esteem Inventory, a non-cognitive measure designed to assess the extent to which children take responsibility for themselves or attribute their successes or failures to the operation of internal or external forces.

Let us now turn to the assessment of the impacts of the curricular programs. Here we are interested in two questions, 1) "How does the progress of Follow Through children compare with that of comparison children?" and, 2) "How does the MAT performance of the Follow Through children compare with that of the national norms group?" Since the Follow Through and comparison children are non-equivalent groups, that is, since the groups differ on a variety of background characteristics including average achievement level at entry to kindergarten, it is necessary to look at the progress that the children have made during their participation in the program rather than compare the raw scores of the treatment and comparison groups. In order to accomplish this estimation of the progress and answer the question,

"How does the progress of the Follow Through child compare with that of the comparison child?" a regression model was developed for each outcome which utilized only comparison children. The regression model involved the pretest score on the Wide Range Achievement Test at entry to kindergarten, family income, occupation, ethnic membership, sex, age at entry to kindergarten, first language of the child, and a variety of other interactions. Since the model utilized only comparison children, it essentially describes progress for that group, that is, given certain background characteristics and a particular WRAT score at entry to school, the model yields a predicted posttest score for each child. The coefficients that make up the model indicate the contribution of each background factor and the pretest to posttest level for the average comparison child. This model that is the set of coefficients was then applied to the Follow Through group. The application of the model to the FT group yielded a predicted score for each child -- that is, the score that the Follow Through child would have received if he had participated in the comparison group experience. The impact of the treatment is assessed by exploring the discrepancy between the predicted score and the actual score that the child obtained. The subtraction yields an effect: a positive effect if the program has had a favorable impact, and the child's score is higher than expected; a zero effect, that is, an effect that is not statistically different from zero if the Follow Through child has achieved or progressed as much in the Follow Through experience as he would have been expected to achieve in the non-Follow Through experience, and a negative effect if the child has an observed score which is less than his predicted score, suggesting that the child might have performed better had he been in the comparison classroom rather than in the Follow Through classroom. These effects are shown in Table 1 for each of the outcome areas. In the table, differences between Follow Through and non-Follow Through are expressed in standard deviation units. Thus a difference of $-.44$ indicates nearly one-half a standard deviation difference favoring non-Follow Through. We restrict our discussion to effects which are both statistically different from zero and which are also as large as one-quarter of a standard deviation which is basically an arbitrary criterion as a device to understand the educational significance of a difference between groups. That is, some differences which may be statistically significant are not in fact of any educational consequence, and one-quarter

of a standard deviation difference provides at least a face valid criterion for educational significance.

The comparison of the Follow Through groups within a particular site with this overall comparison model has several advantages. First, the regression coefficients, developed using only the comparison children, are not affected by the treatment. If a Follow Through program produced greater or smaller gains in some subgroup of the children with which the program was dealing, for example, if children with higher achievement scores in Follow Through gained more, and if the coefficient were developed as for an analysis of covariance using all children, the coefficient representing the contribution of a background measure to the outcome measure would be biased by treatment. A second advantage of the use of the overall comparison model is that it provides a common comparison for all programs and for all sites.

The second method of assessing the impact of the program is found in the assessment of the norm scores of the children at the end of third grade. Although it is difficult or impossible to attribute to a program the normative status of a group at the end of third grade, it is certainly not expected that these groups would be performing at or near grade level since the majority of children are "disadvantaged." Furthermore, the legal definition of educational disadvantage includes a criterion concerning the performance of children in terms of grade level.* The legal definition of disadvantage defines students who are performing one or more years behind the expected grade level as educationally disadvantaged. This then provides a natural criterion for exploring the status of groups at the end of third grade. Groups where the median score is at or near grade level cannot be construed as disadvantaged, at least not legally, and therefore programs or sites in which the median score is close to grade level can be construed as successful. The grade level performance of the children in each of the sites associated with the program are shown in charts such as Figure 3. The figures show the median grade level for MAT Reading, Math, and Spelling for each site and the percentile of the median.

You may ask, "Why are these people proposing the use of grade equivalent scores which have been in disrepute for a long time in the field

*Section 17a of General Provisions Act as amended by public law 93-308 defines educational disadvantage as achievement which is one or more years behind that expected at the appropriate grade level.

of educational evaluation?" This is indeed a legitimate question since the opportunity to attribute the level of performance of the children within a given group is minimal or absent when only grade equivalent scores at the end of third grade are available. Still, these scores have face validity in terms of their use in schools. Teachers, not all, but some, utilize norm scores to make judgments about the assignment of children to special programs. Furthermore, norm scores are often used to communicate a child's status and progress to parents. The statement that "Tommy is reading a grade level ahead or behind" has great significance for Tommy, his parents, and his teacher. Furthermore, we probably all agree that the preferred estimate of program impact, that is, the regression estimate which is provided by comparing the observed and predicted scores of the children at the end of third grade is also suspect and subject to a variety of criticisms. The majority of these criticisms suggest that regression approaches underestimate treatment effects. That is, they show no effect on occasion, or negative effects when in fact the program has had a favorable impact. It is beyond the scope of this presentation to go into these problems in detail; however, it would appear that both grade equivalents and regression estimates of progress have a variety of distinct advantages and disadvantages. It makes sense to capitalize on the information available and combine these two information sources about programs and the status of children and define a favorable program impact in a somewhat unorthodox way. Basically, a favorable program impact can be defined either as greater progress in Follow Through than expected or as progress in Follow Through which is no less than expected, paired with performance on the norms which are at or near grade level. This definition of program impact is utilized in the following presentations.

Let us now turn to some of the limitations of this study and of the data to be presented in this symposium. Indeed, we could devote the entire symposium to an exploration of the limitations of the Follow Through evaluation. There are various design constraints, aspects of the measurement battery, implementation issues, attrition issues, analysis issues, and others which we could consider in detail. However, one major limitation seems to stand out among the problems. This is the issue of Head Start and preschool, and the possible consequence of not utilizing this aspect of children's background in the model of scores. First of all, I must point

out that this aspect of the children's background was not selectively ignored. The variable was not utilized in part because it is severely confounded with site and Follow Through. In some sites 95% of the Follow Through children may have preschool or Head Start experience vs only 30% of the comparison group; thus, this variable does not make for a useful covariable. Furthermore, the missing information rate -- that is, information as to whether a child participated in preschool -- is intolerably high in many cases. Since Follow Through was intended as a follow through on Head Start, the absence of this variable is unfortunate. Furthermore, since preschool may affect the pretest score, the Wide Range Achievement Test, at entry to kindergarten, may further bias against the finding of treatment effects since there generally appears to be more preschool graduates in Follow Through.

COGNITIVE-DISCOVERY CURRICULAR MODELS IN PROJECT FOLLOW THROUGH

John C. Larson, Ph.D.

Abt Associates Inc.
55 Wheeler Street
Cambridge, Mass. 02138

Overview of the Cognitive-Developmental Models

Two of the major curricular models in Project Follow Through exemplify a form of the cognitive-discovery approach to early education. They seek to stimulate the child's basic cognitive abilities and problem-solving skills as the foundation for later academic achievement. In addition, they attempt to enhance the child's self-esteem and feelings of competence in support of the child's curiosity, self-initiated exploration of the learning environment, and self-monitoring strategies for solving problems. The sponsors of both these models see active parent involvement in the child's learning at home and in class as an essential component in the child's adjustment to the school and to learning in general.

While both approaches focus on the child's development of problem-solving skills, Sponsor A emphasizes learning experiences structured around a theoretical framework of child development in classification, seriation, and spatial and temporal relations, while Sponsor B emphasizes a more open responsiveness to a range of child-directed learning interests. Both sponsors stimulate self-initiated and self-monitored learning activities in children, but Sponsor A stresses the child's prior planning and follow-up more than does Sponsor B. Sponsor B supports self-confidence by more consciously stressing mutual respect for subcultural differences and responding to the child's own individual range of interests. Sponsor A supports self-esteem by more directly encouraging self-expression of thoughts, plans and learning experiences in the written language. The approach of these sponsors differs from the Behavioral approach in their different concepts of child development and in their lack of specific emphasis on academic skills. While their objectives are similar in many respects to the Psychodynamic approach, their relatively greater emphasis on problem-solving skills over socio-emotional and interpersonal skills differentiates them from the latter.

Sample

The sites for the analysis of Sponsor A include two large cities and a southern community. The children in these three sites come from some of the most economically disadvantaged families in the total Follow Through sample. With a mean annual income of \$3703, these sites reported some of the lowest income levels and highest proportions of single-parent families. Their kindergarten entry scores were among the lowest in the total analytic sample. Fewer than half of the mothers had completed high school.

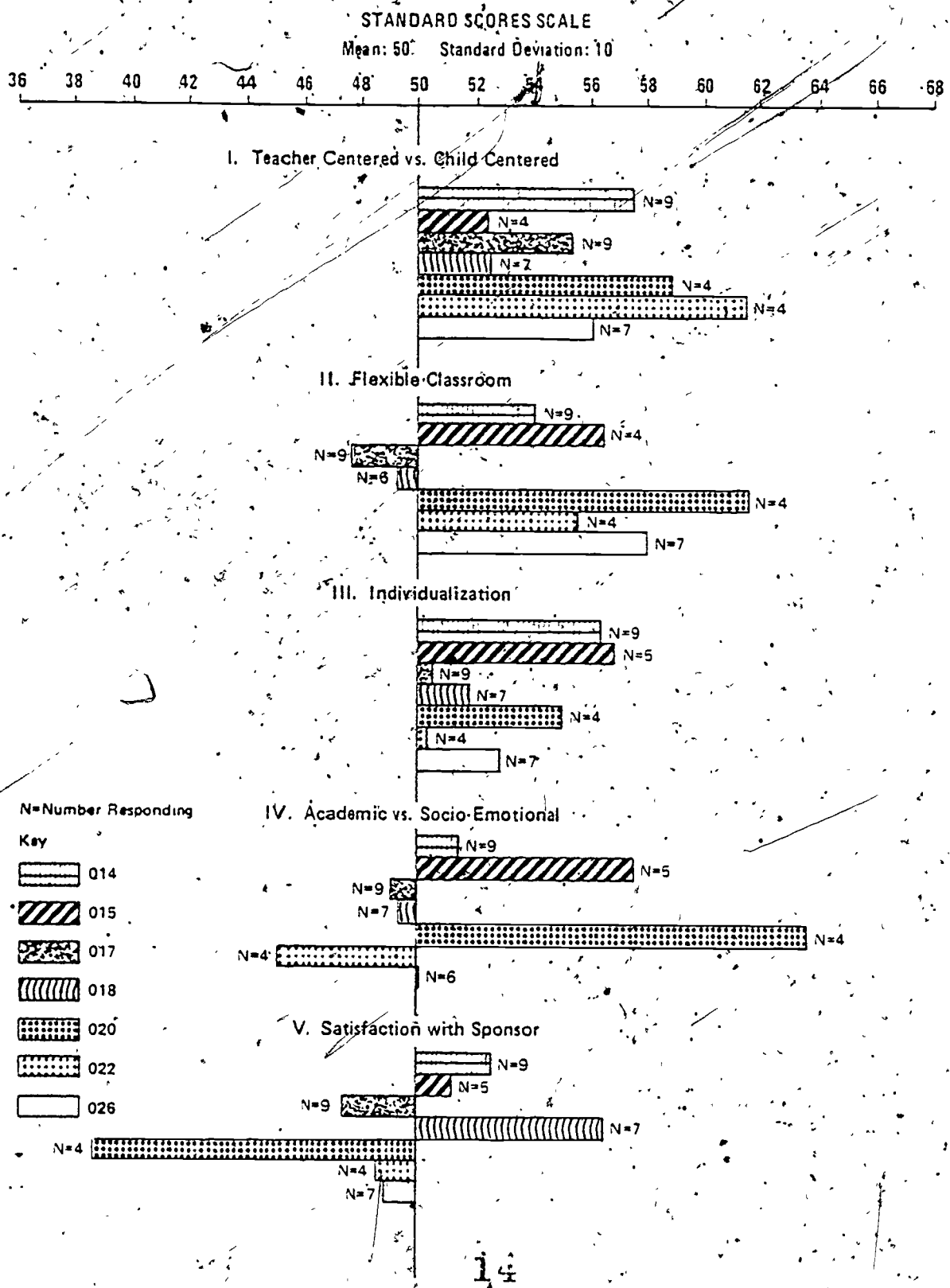
The sites included in the analysis of Sponsor B had generally higher levels of annual income (with a mean of \$5756) than those found in most other Follow Through sites. A greater proportion of mothers at these sites had completed high school than was true for the Follow Through average, and the children scored higher on academic tests at kindergarten entry than the mean of the overall analytic sample. Thus, the environmental circumstances in which this sponsor implemented the program were not so educationally handicapping as those of Sponsor A.

Teacher Reports as Indicators of Program Implementation

The data in Figures 1 and 2 illustrate the extent to which the teachers in the separate sites of these programs characterized their classrooms as: 1) Child-centered vs teacher-centered; 2) Flexibly structured vs inflexibly structured; 3) Individualized vs group oriented; 4) Socio-emotional vs academically oriented; and 5) Teacher's satisfaction with the model.

These data suggest that the cognitive-developmental models generally succeeded in implementing a child-centered, flexible, individually-oriented curriculum with a relatively greater emphasis on socio-emotional than academic objectives. Although Sponsor A has only three sites included in the analysis, its classrooms scored consistently in the same direction on all of the teacher reports. This uniformity of teacher reports, found in no other sponsor discussed in this symposium, suggests that Sponsor A had a well-integrated training program with respect to the dimensions reported here. For Sponsor B the sites were unanimously more child-centered and individualized than the average of all teachers in the analytic sample. There was, however, evidence of site variability in the Emotional-Academic Teacher Satisfaction, and Flexibility dimensions. Given the complexity of

Figure 2
Indications of Implementation From Teacher Questionnaire Clusters
for Sponsor B
Cognitive Development Approach



the Follow Through models, we cannot assume that all meaningful curricular differences are registered on just these dimensions, nor can we assume that teacher reports directly correspond to the levels of implementation desired or specified by sponsors. These data do indicate, however, that characteristics important to a cognitive-developmental approach were successfully communicated to the minds and understanding of teachers.

Interpretation of Curricular Impact on Child Outcomes

Considering next the child outcomes at the end of third grade, we examine curricular effects from three sets of data: first, the comparison of the Follow Through groups to the pooled non-Follow Through group; second, the contrast of the Follow Through groups to national achievement norms; and finally, the qualitative contrast of curricular effects across different domains.

Levels of Curricular Impact

The data in Figures 3 and 4 suggest how the performance levels of the Follow Through groups differ from those expected of similar groups in non-Follow Through. In these figures the bars extending to the right of the center suggest that the Follow Through site performed at a level higher than that predicted by the regression model, while bars extending to the left suggest the converse. In the academic areas, the effects suggest that the children in Sponsor A performed somewhat lower than would be expected. Of the three sites, the one which indicated higher than expected levels of performance on Spelling, self-esteem, and conceptual problem solving relative to the other two sites, is that site in which the teachers also reported higher levels of satisfaction, greater classroom flexibility, and the least teacher-centeredness. The classroom emphasis also tended quite strongly toward the socio-emotional rather than the academic direction.

The academic effects for Sponsor B are somewhat more mixed across sites than for Sponsor A. Two of the seven sites indicate small but consistently better than predicted performance on Math and Reading, while four of the seven sites indicate performance levels slightly to markedly lower

Figure 3
 Sponsor Effects on Seven Third Grade
 Tests, Expressed in Standard Deviation
 Units of Raw Score, For Sponsor A
 Cognitive-Development Approach

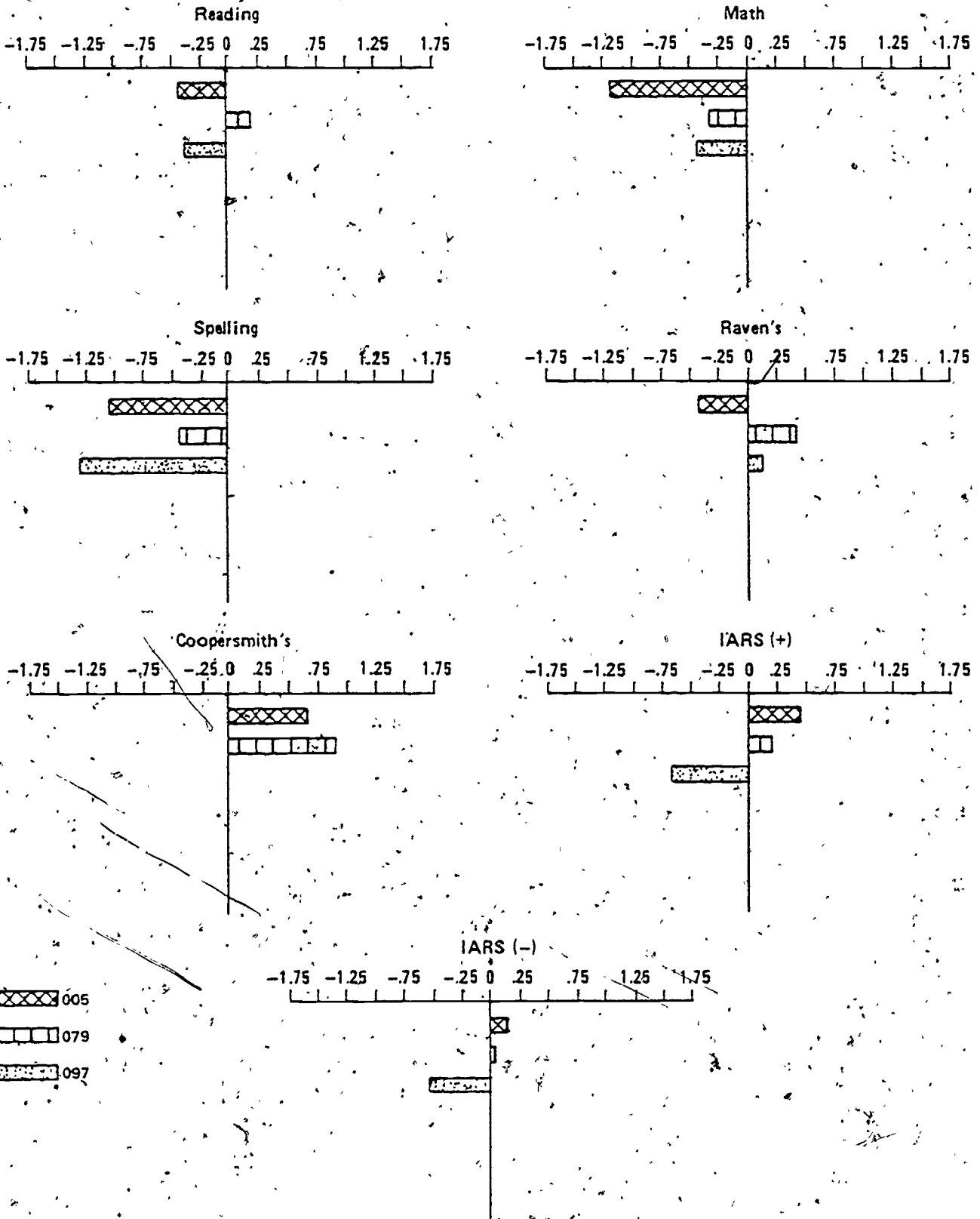
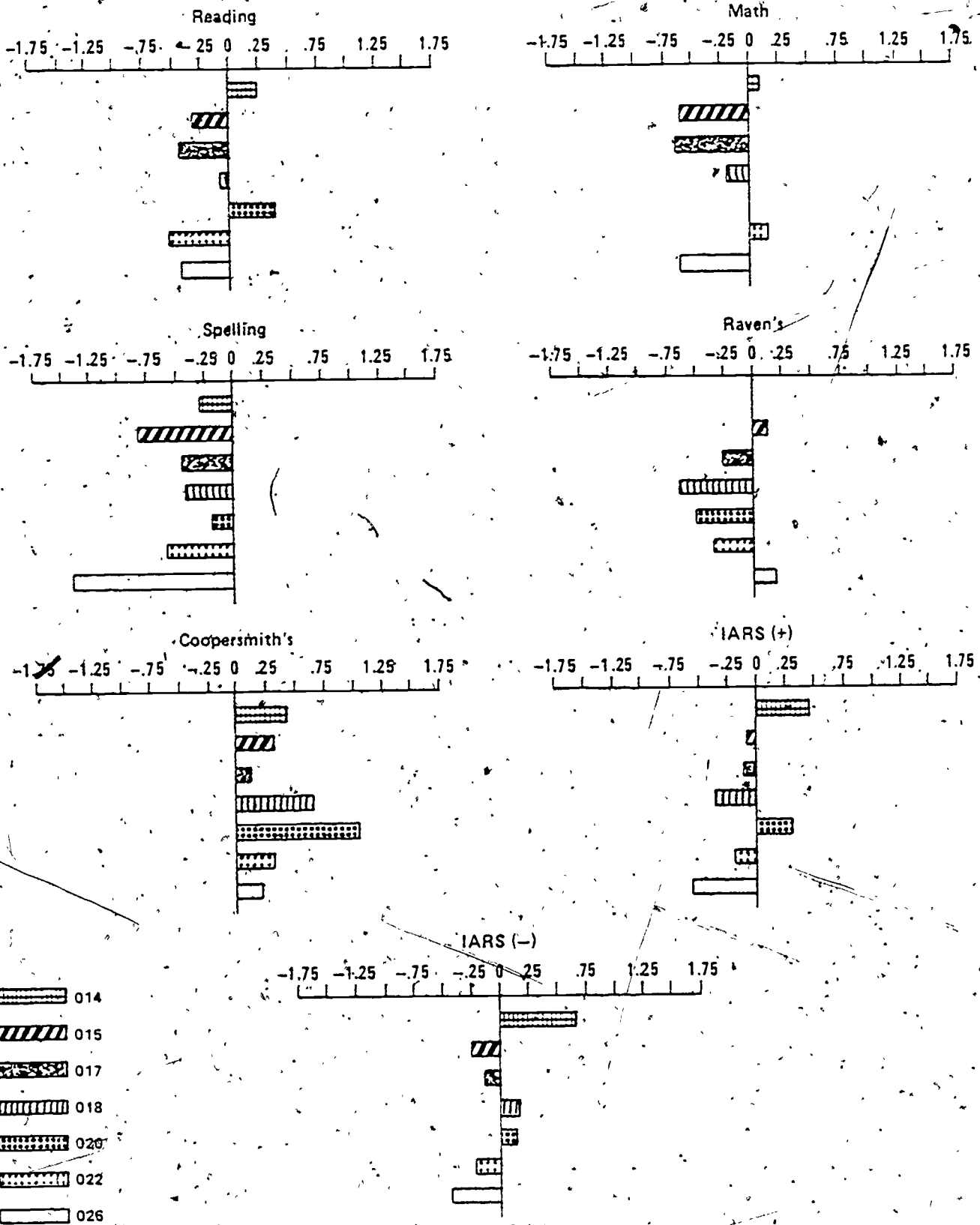


Figure 4
 Sponsor Effects on Seven Third Grade
 Tests, Expressed in Standard Deviation
 Units of Raw Score, For Sponsor B,
 Cognitive-Development Approach



17

than predicted by the regression analysis. All sites scored lower than expected in Spelling. As was the case for Sponsor A, the teachers in the sites with relatively higher levels of effect tended to report the highest levels of satisfaction and relatively high levels of socio-emotional emphasis, non-teacher-centeredness and flexibility in the classroom organization.

It is difficult with these data to draw too close a correspondence between levels of teacher reports and levels of curricular effects, because of both imprecision of the measures and lack of consistency between these domains on the low end of the scores. However, within the range of site variability for both sponsors, the most positive academic effects seem to follow those sites wherein the dynamics of the cognitive-developmental approach are most in evidence.

In the affective domain and conceptual problem-solving test, the effects of Sponsor A generally parallel those observed in the academic domain. Two of the three sites tended to perform at lower levels than would be expected, while the third site tended to perform at higher levels than predicted by the regression equations. In two sites there is some indication that where lower academic performance is found, affective scores are not so low, or where higher than expected academic scores are found, the affective scores are still higher. For Sponsor B the affective and problem-solving areas indicate a pattern of site variability similar to that found in the academic areas. As noted above, in sites where children score lower than expected in academics, they seem to perform relatively better on the affective tests.

To summarize the comparison of the cognitive-developmental curricula to the pooled non-Follow Through, the clearest generalization seems to be site variability. This analysis suggests that while some sites show higher than expected performance in several areas, most of the sites for this approach tend to show performance levels somewhat lower than would be expected from the regression analysis. There is some evidence to suggest that relatively higher levels of performance appear in those sites reporting a higher degree of classroom flexibility, child-centeredness and teacher satisfaction.

While the results of the regression analysis provide comparisons of the programs to the pooled non-Follow Through group, grade equivalent and standardized scores on the Metropolitan Achievement Tests in Figures 5 and 6 provide a comparison of the models to another external criterion, the

Figure 5
Grade Level Equivalent and Standardized Scores on MAT Subtests
(Reading, Math, and Spelling) with Income Levels for
COHORT II Sites in Sponsor A
Cognitive – Developmental Approach

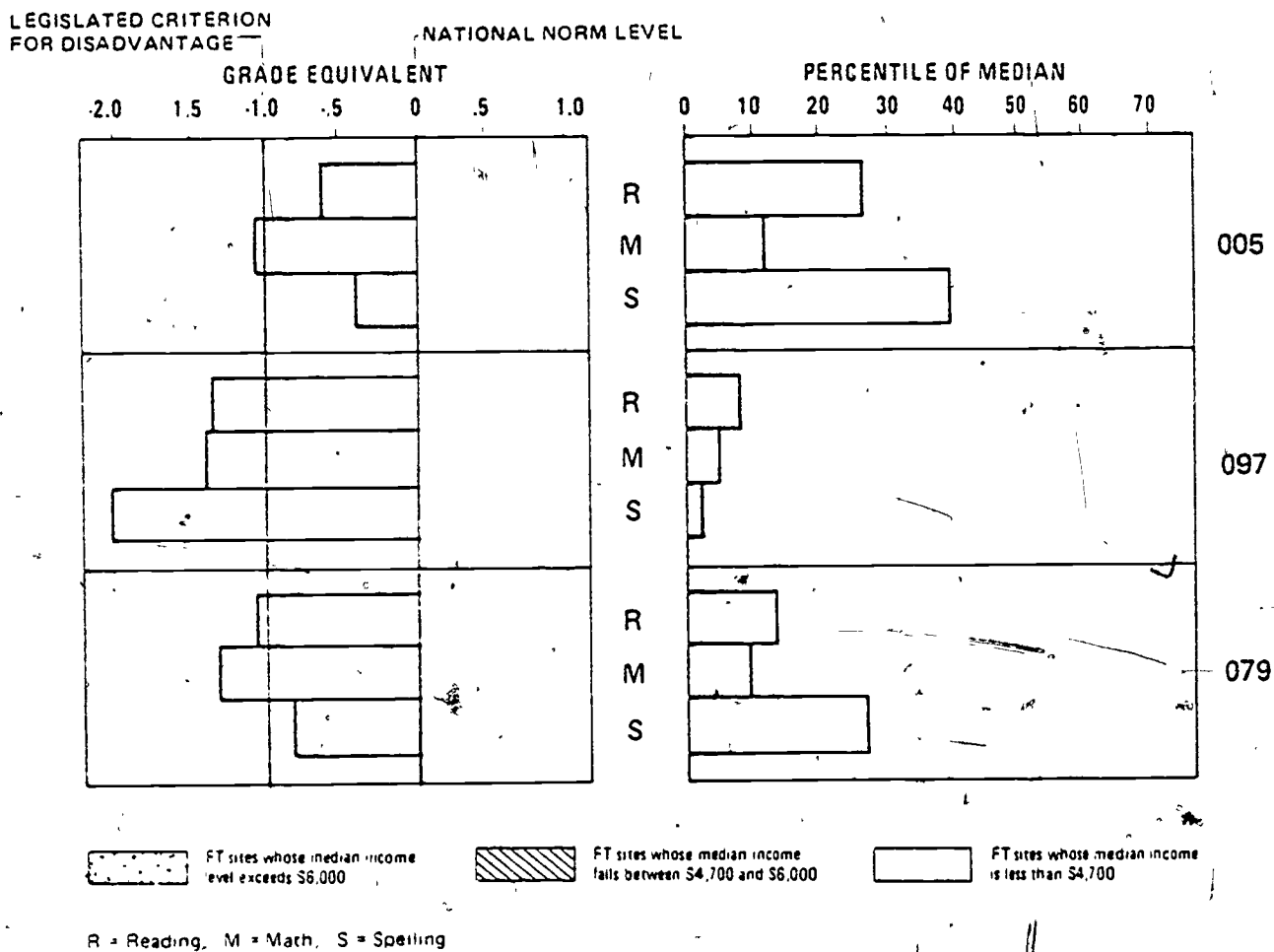


Figure 6
 Grade Level Equivalents and Standardized Scores on MAT Subtests
 (Reading, Math, and Spelling) with Income Levels for
 COHORT II Sites in Sponsor B
 Cognitive - Developmental Approach

LEGISLATED CRITERION
 FOR DISADVANTAGE

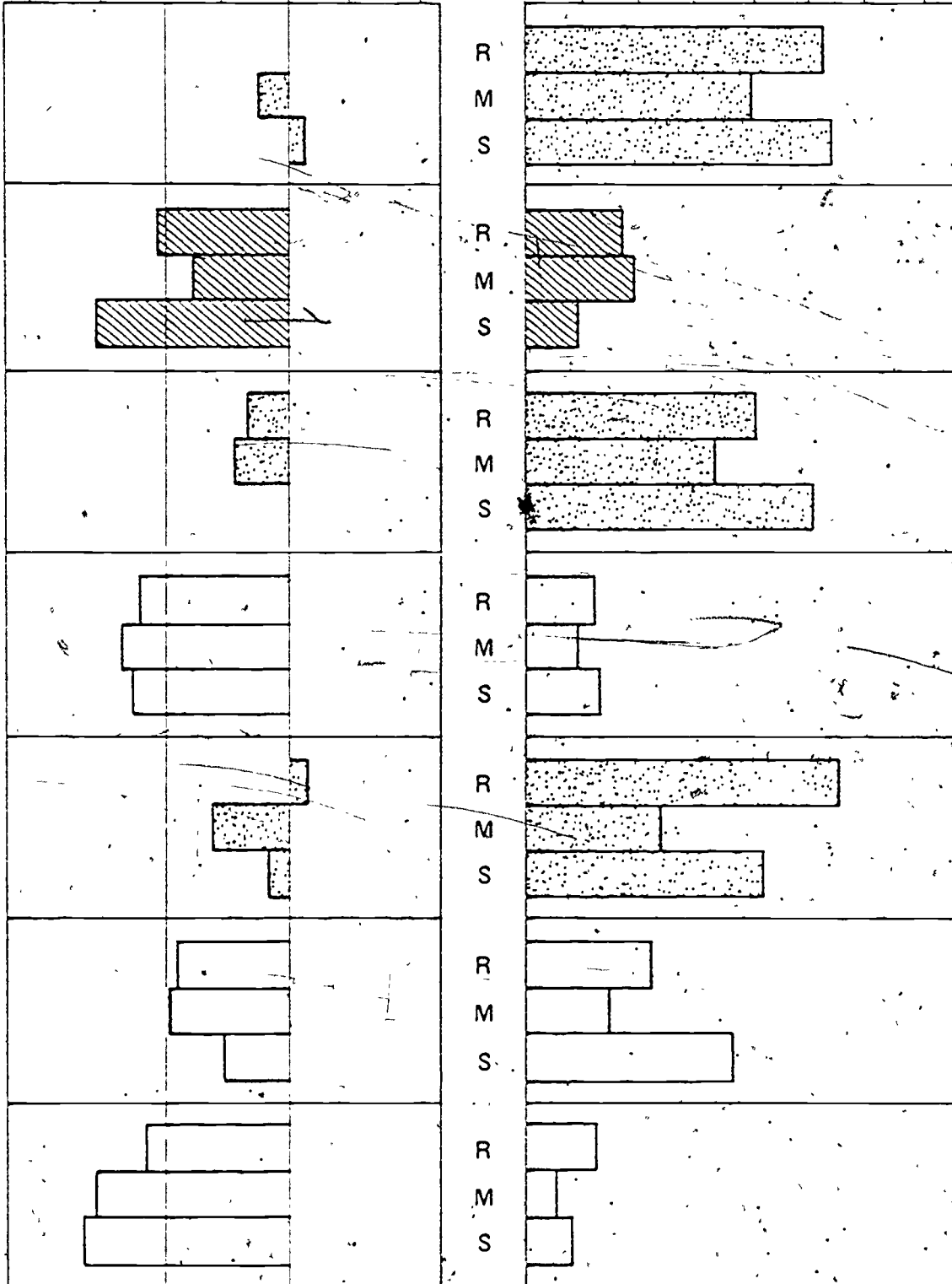
NATIONAL NORM LEVEL

GRADE EQUIVALENT

PERCENTILE OF MEDIAN

-2.0 -1.5 1.0 .5 0 .5 1.0

0 10 20 30 40 50 60 70



FT sites whose median income level exceeds \$6,000

FT sites whose median income falls between \$4,000 and \$6,000

FT sites whose median income is less than \$4,000

R = Reading, M = Math, S = Spelling

national norms for these tests. The comparisons just discussed suggest levels of program impact, while the grade equivalent scores suggest the status of the Follow Through sites relative to the country at large by the end of the children's tenure in the program. For Sponsor A the academic status of the three sites can be rank-ordered as: 1) within grade-level performance, 2) borderline grade level performance, and 3) below grade level performance. This order does not correspond directly with the sponsor effects analysis, since the higher-than-expected effects were found in the borderline grade level site. The data suggest that the order of grade level performance does correspond, however, with the rank of the sites' mean income levels.

For Sponsor B we find a similar relationship between grade level performance and site mean income. Four of the sites are within the grade-level performance criterion, and three sites are below grade-level performance criterion. What is the relationship of the grade equivalence scores to the program effects discussed above? One hypothesis would suggest that program effects are more related to levels of implementation, while grade equivalent status is related more to income level. As a corollary, it may be that sites with grade equivalence levels higher than that expected from their income levels obtained these achievement levels in part through a positive impact from the sponsor's program.

Qualitative Assessment of Curricular Impact

Analysis of the levels of impact for the curricular model seeks to answer the question, "How much impact did a curriculum have on child outcomes?" An appropriate answer to this question is best obtained through a rigorous experimental context, yet the metal of our experimental logic inevitably loses much of its temper when subjected to the pressures and strains of the real-world environment of teachers and children scattered across the entire country. In a quasi-experimental setting, the results of our analyses are sensitive to a number of circumstances such as the degree of similarity between the Follow Through and comparison groups, the absolute level of poverty with which the local community is burdened and in which the sponsor seeks to bring about change, the specificity of the instructional program to the test instruments, and other contextual features of the community.

A fuller consideration of curricular impact, however, demands consideration of still another set of factors. Within the constraints of any one site, what was the quality of the program's impact? Considering the content areas relative to each other, in what domains were the program's effects most positive (or least negative) compared to those areas which tended to be the last to respond to the curricular influences? The factors affecting these issues refer to the dynamics of the learning situation, or the form of educational experience molded by the program. To what extent are learning activities self-directed or teacher controlled? What is the scope of generalization and flexibility of application to which the learning is applied? Does the learner acquire self-esteem through taking risks and discovering competencies, or through following directions well, or by learning to enjoy learning itself and mutual social support? These issues, too, must be examined in assessing a curricular program's influence.

As a heuristic approach to this question we have rank-ordered the levels of effect across the seven outcome measures within each site. That outcome which showed the highest positive (or least negative) effect, regardless of its level, was given the highest rank, and vice versa. We then determined the mean rank for each content area across the sites of the cognitive-developmental approach. The results, indicated in Table 1, suggest that for the cognitive-discovery approach, relatively more positive effects were obtained in the area of self-esteem, followed next by internal locus of control and problem-solving. Conversely, Math and Spelling were the last areas to show curricular effects. The preeminence of the affective areas over the academic areas corresponds to the relatively greater teacher emphasis reported on the factors of child-centered, socio-emotional focus, and classroom flexibility for this curricular model. Based on the goals of the approach, however, we would have expected to see the problem-solving area emerge as that which is first affected positively by the curriculum. These data suggest that the dynamics of the cognitive-developmental curriculum correspond in part with the aspirations of the model. Further studies designed specifically to examine such dynamics, and to determine the relationship of such learning orientations to future school success must be conducted in order to gain a fuller understanding of the ultimate utility of the cognitive-developmental curriculum in preparing the child for later learning demands.

Table 1

Means of Content Area
Ranks Across Models

Model	Reading	Math	Spelling	Raven's	Coopersmith	IARS (+)	IARS (-)
Behavioral	4.2	4.2	1.3	2.6	6.3	4.5	4.7
Cognitive-Discovery	3.8	2.8	1.5	4.0	6.8	4.5	4.6
Psychodynamic	3.6	3.7	1.0	5.2	6.7	4.4	3.3

BEHAVIORALLY STRUCTURED PROGRAMS IN PROJECT FOLLOW THROUGH

Linda B. Stebbins, Ph.D.

Abt Associates, Inc.
55 Wheeler Street
Cambridge, Ma. 02138

Introduction

Follow Through as an experimental program was implemented using a strategy known as "planned variation." This strategy systematically introduced a variety of well-defined programs into the kindergarten through third grade years of public education in certain school districts throughout the United States. To operationalize the concept of "planned variation," USOE involved educational specialists, each sponsoring a different educational approach in a variety of school districts. The Follow Through program as it was defined by these matched pairs of models and school districts consisted of two sets of components: social service and education. One set of components -- social services, parent, and community participation -- was mandated in the original enabling Congressional legislation. The second component, education, was the essence of planned variation. This component was comprised of four major types of educational strategies:

- Projects emphasizing a behaviorally structured approach to acquire academic skills, particularly reading and arithmetic,
- Projects stressing cognitive thinking through asking and answering questions, problem solving, and creative writing,
- Projects emphasizing social emotional development and encouraging exploration and discovery in academic areas, and
- Projects focusing on preparing parents to improve the education and development of their children (GAO, 1975, pp. 3-4).

This paper is designed to explore the measured outcomes of one of these educational strategies -- the behaviorally structured approach.

Goals of the Behaviorally Structured Models

There are two models which can be clearly identified as prescribing a behaviorally structured approach. Drawing upon behaviorist theories of children's actions as an interrelated system of stimuli and responses, these two models have developed curriculum approaches which focus on the assumption that children's behavior is primarily a result of external stimuli. By carefully controlling the external stimuli, these models attempt to modify the children's behavior in ways which meet the specific goals of the program. Meeting goals is accomplished by clearly defining a series of discrete steps, and working toward success at each step. Praise and rewards are provided for successful completion of each step.

There are some specific differences between these two programs. First, let's label the models C and D (to distinguish them from other models discussed in the symposium) and then examine those differences. Model C's goals focus on accelerating the learning of reading, math, and language skills through the use of programmed instruction. (Non-core subjects are generally introduced after mastery of basic skills, usually toward the end of third grade.) The model is operationalized through the use of sponsor-developed programmed instructional materials in small homogeneous groups in sound-controlled booths. The teachers present a fast-moving series of specific questions to elicit verbal student responses. Proper student responses are reinforced and wrong answers are corrected according to specified procedures. Student groups rotate by schedule. Group instruction is accompanied by individual, self-directed practice in workbooks. Planned homework assignments are also coordinated with the classroom lesson. Criterion reference tests administered to students at frequent and regular intervals provide information to the teachers on the students' progress.

Parents participate in the program in several capacities. Some are employed in each classroom on a permanent basis as teacher aides (two per classroom) and assistants. Others are employed as needed and trained to administer the criterion-referenced pupil progress tests and operate the videotape equipment to film the teacher at work in the classroom. Still others are employed as family workers. In this capacity,

they acquaint parents with the Model C program, provide specially developed materials which parents can use at home to supplement classroom instruction, make available to those parents who so desire a sponsor-developed programmed course in child management, encourage participation in PAC meetings, and assist in training the classroom aides and assistants. In sum, this model is designed to help children achieve quantifiable objectives in the basic skills of reading, arithmetic, and language through the application of behaviorist principles to instructional methodologies and materials.

Like Model C, Model D is highly structured, derived from behavioral psychology, and designed to increase students' reading and arithmetic skills. In addition, Model D includes handwriting and spelling skills as core subjects, and emphasizes the development of social skills as well. Model D is based on the premise that the conscious and systematic use of praise and approval will promote academic and behavioral growth in the individual child. The model is operationalized by establishing a "token economy" within each classroom. Teachers award tokens for approved social and academic performance. The children can use these tokens during an exchange period to purchase activities of their choice, such as games, toys, books, etc. Tokens and praise are distributed according to individual rates of progress rather than for group performance. Teachers in this model may choose among either sponsor-developed or commercial learning materials, but they are encouraged to select those which can be adapted to the model. Using a machine-readable data form, teachers prepare continuous progress reports on students. The data is then computer analyzed and an individual progress prescription is returned within a day.

Parents also play an important role in the Model D classrooms. Two parents are trained and placed in each classroom for an eight-week period to teach handwriting and spelling, to do individual tutoring, and to become knowledgeable in the use of the Model D techniques. At the end of the eight weeks, the teaching parents may continue or not as they choose. Although many parents serve only for an eight-week session and teach only in one curriculum area, some teach a full year in as many as three curriculum areas. Many eventually become permanent teachers' aides.

In sum, acquisition of both academic and social skills are encouraged through the systematic use of positive reinforcement in the form of praise and tokens.

Teacher Reports as Indicators of Program Implementation

Determination of whether or not a particular planned variation strategy is effective requires both a clear description of the essential elements of the educational program and some indication of whether or not implementation has been attained. In the preceding section of this paper the goals of both models are presented with an explanation of the elements of the learning environment, both within the classroom and at home. Identification of evidence of implementation, however, is not readily available. Recognizing that there are numerous conditions which could influence the implementation of the models, it still seems reasonable that after a minimum of five years operation there should be some indications of implementation of the intended classroom environment which could be identified from the data in the teacher questionnaire. The teacher questionnaire is the only measure which provides us with data across all sites included in the analytic sample and does reflect the perceptions of teachers who operationalize the models daily in the classroom. While we recognize that these are data on reported rather than observed practices, we nevertheless feel the information is valuable. Thus, in order to assess implementation, or some proxy for implementation, we have selected specific items from the teacher questionnaire administered to all Follow Through and non-Follow Through teachers at the end of third grade.

Figures 7 and 8 display five dimensions of the classroom learning environment which can serve as proxy measures of the model's intended goals; teacher centeredness; flexible classroom; individualization of instruction; socio-emotional versus academic orientation; and teacher satisfaction with the model. Referring specifically to Figure 7, Model C has generally implemented a structured classroom learning environment as intended; however, there are some notable site variations in the degree of classroom structure. Site 005 and 081 have clearly implemented highly structured classroom learning environments whereas Sites 069 and 072 are much less structured. All sites show below average individualization

Figure 7
Indications of Implementation From Teacher Questionnaire Clusters
for Sponsor C
Behaviorally Structured Approach

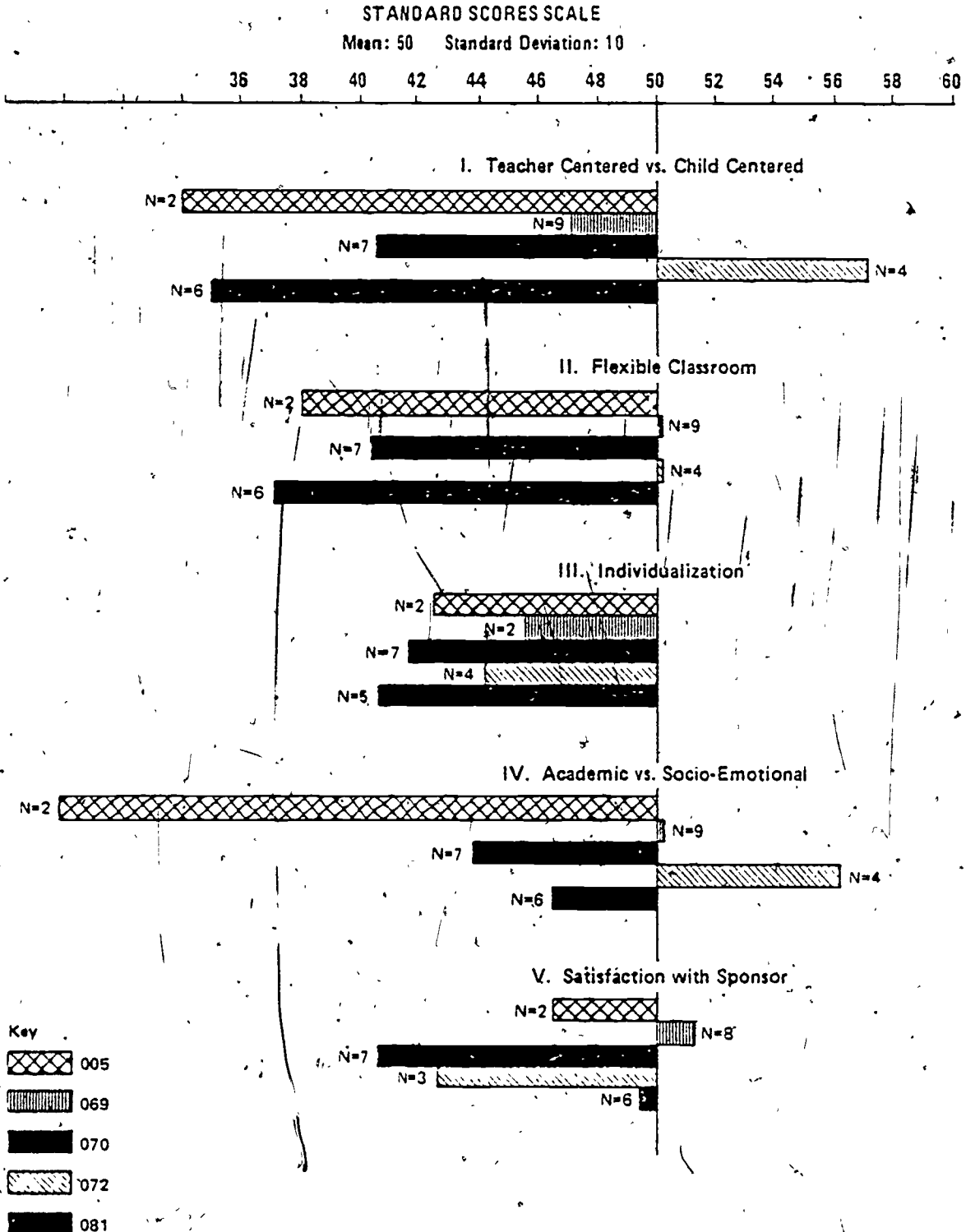
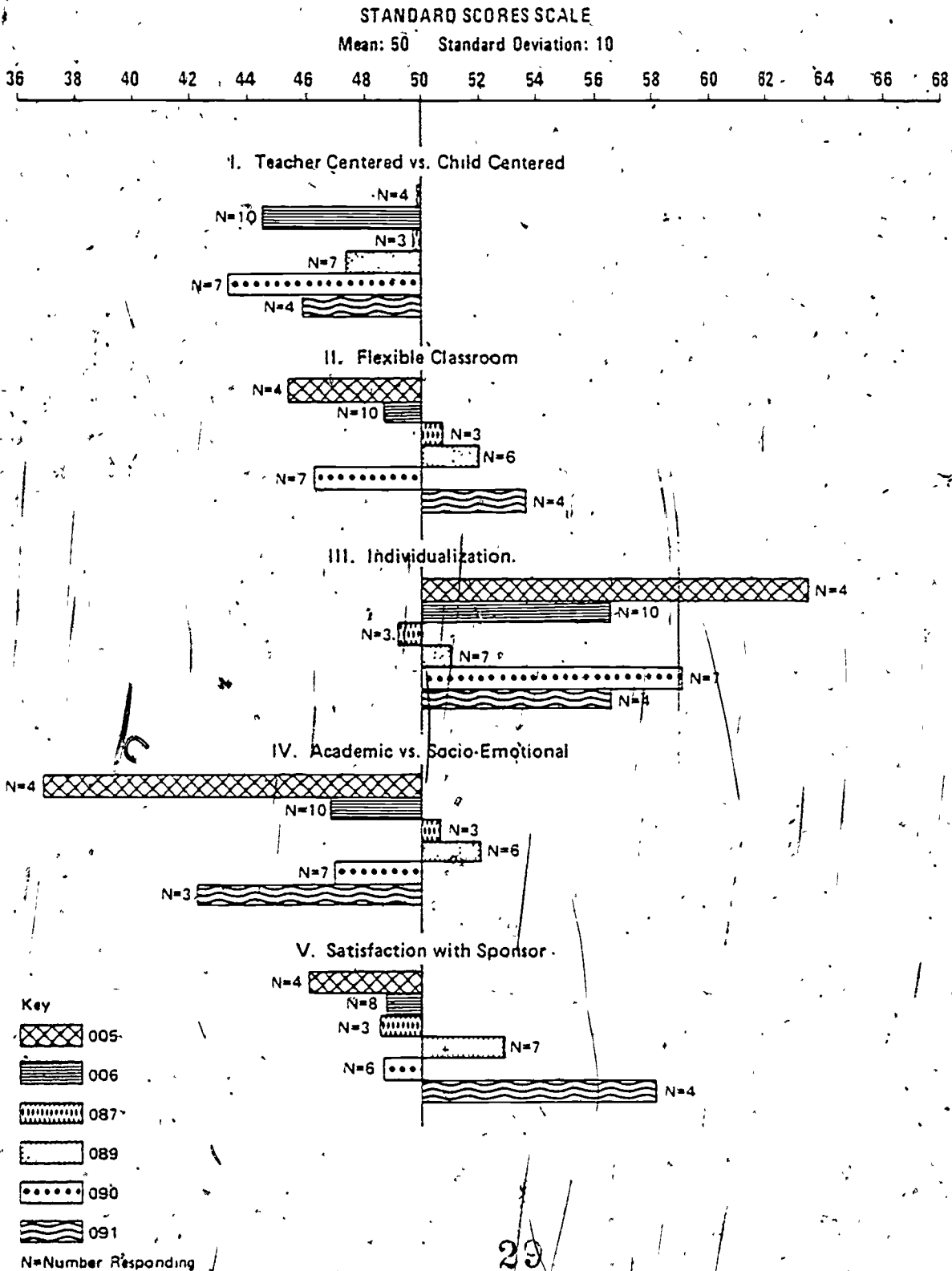


Figure 8
 Indications of Implementation From Teacher Questionnaire Clusters
 for Sponsor D
 Behaviorally Structured Approach



of instruction. Three sites, 005, 070, and 081, are very high in the degree of academic emphasis. This appears to indicate that the desired degree of academic emphasis has not been achieved in at least two sites, especially 072. Teachers in three sites, 005, 072, and 070, report a low degree of satisfaction with the sponsor. In summary, it appears that the Model C is fairly well implemented, except for site 072, which is deviant.

Figure 8 shows that, generally speaking, Model D has implemented a structured classroom learning environment as intended. Model D sites also show a high degree of individualization of instruction which is consistent with the model. In this case, there is considerably more site variation in the degree of structure implemented. Sites 087 and 089 appear to have less structured programs than other Model D sites. Sites 087 and 089 are also considerably lower than others in the degree of individualization and academic emphasis. Site 091 appears to have adopted a more flexible classroom than might have been expected from the model's description. With the exception of sites 091 and 089, Model D teachers express less than average satisfaction with the sponsor.

In summary, both Models C and D appear to be implemented in most sites.

Sample

The analyses for Model C are based on data from five sites, including one large northeastern city and four cities in the midwest. The Follow Through children in these sites are very similar to the average Follow Through child with respect to such variables as preschool attendance (89% attended preschool), entering pretest scores (33 points on WRAT), child's ethnicity (88% Black children), and family income (\$5200/yr.). The same holds true for the non-Follow Through children included in the analyses of these sites. That is, 47% attended preschool; children scored 33 points on WRAT at entry; 80% are Black children; and the average family income is \$5308 -- all similar to average for NFT children. In general while there is variability across sites, the model average for both FT and NFT groups is similar to the respective average for all Follow Through sites.

Model D, sampled in six sites, has a considerably lower average family income, that is, \$3890 with a range between \$2075 and \$4906. The average for the Follow Through children in Model D's sites is also about \$1000 lower than the average for the non-Follow Through children in these sites. Other variables such as pretest scores, percent with preschool, and percent of mothers with high school education are similar to the average for all Follow Through sites. While the Follow Through children in Model D's sites tend to have a slightly higher pretest score in comparison to the non-Follow Through children, about 20% more of these children have preschool experience.

Interpretation of Curricula Impact on Child Outcomes

To explore the question "Do the models result in impacts either prescribed in their stated goals and objectives or reflected in other areas measured by the test battery?" let us use two comparisons of data. One, a comparison of the observed performance of the Follow Through children with the performance expected in the absence of treatment. And two, the comparison of the Follow Through children with the national norms group for the Metropolitan Achievement Test.

Using the difference between observed and predicted scores for Follow Through children (revealed by the pooled comparison analysis), sites can be categorized in three ways: 1) Follow Through performed better than would be predicted without the program; 2) Follow Through performed as well as would be predicted; 3) Follow Through performed less well than would be predicted. As Figures 9 and 10 show, these categories are distinguished in terms of standard deviation units as follows: 1) if Follow Through performs at plus a quarter of a standard deviation unit better than non-Follow Through, then we may say Follow Through performed better than would be predicted without treatment; 2) if the number is negative and greater than .25, we say Follow Through performs less well than would be predicted without the program.

First, let us examine the area of basic skills, as measured by the Reading, Math, and Spelling outcomes. Looking at Figures 9 and 10 it can be seen that there is considerable variability both across the sites within each model and within each site across outcomes. Both

Figure 9
 Sponsor Effects on Seven Third Grade
 Tests, Expressed in Standard Deviation
 Units of Raw Score, For Sponsor C
 Behaviorally Structured Approach

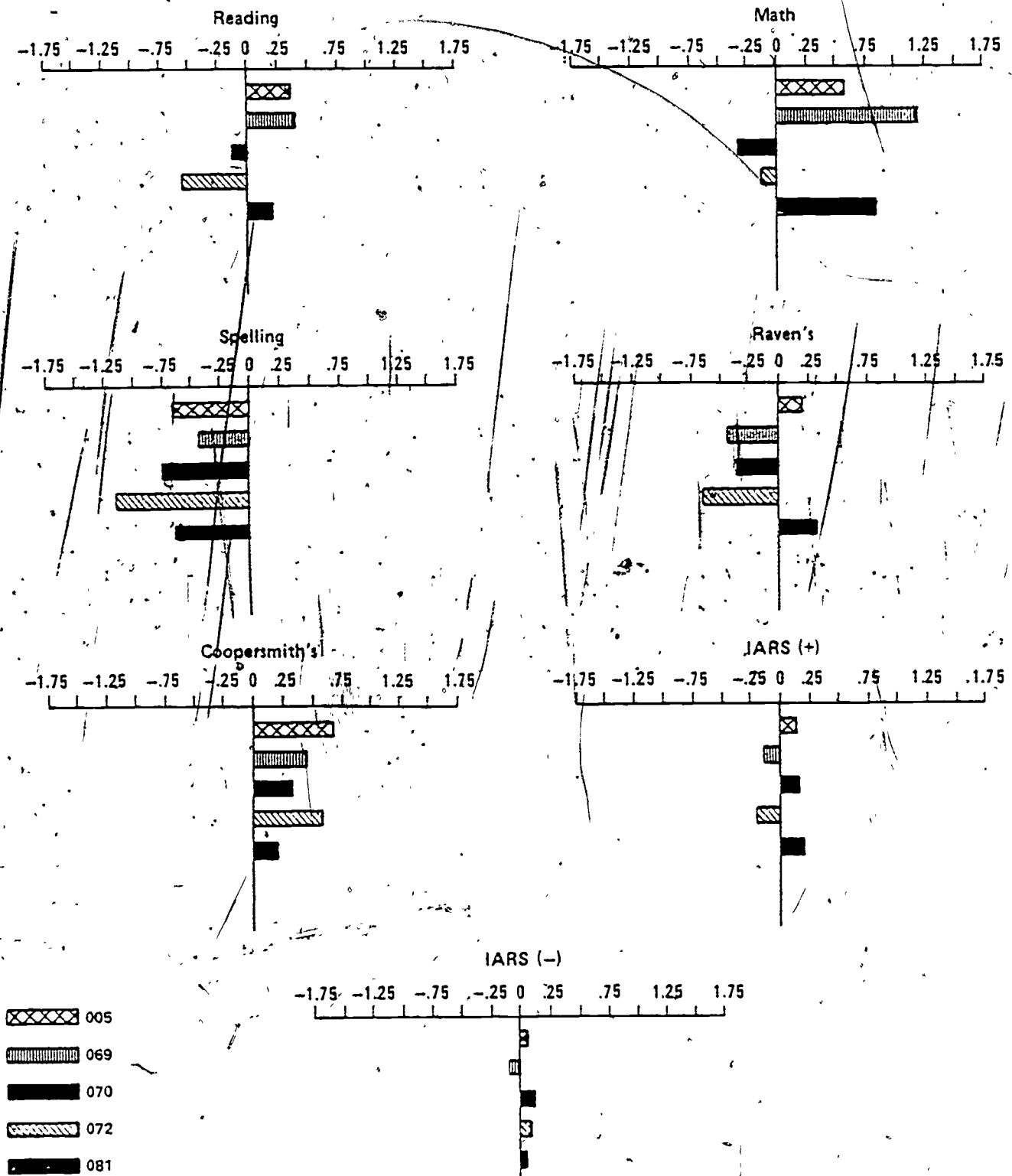
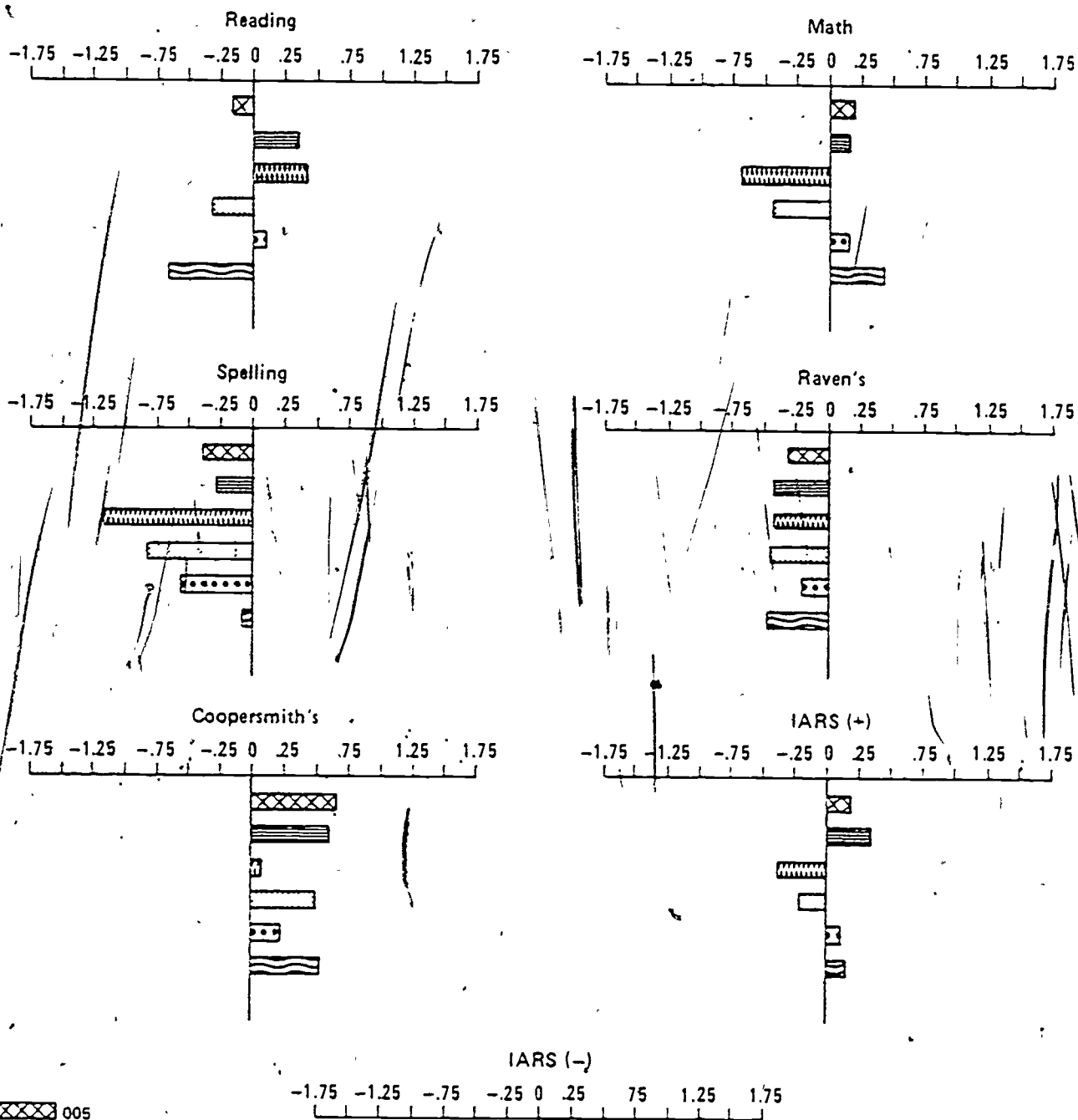


Figure 10

Sponsor Effects on Seven Third Grade Tests, Expressed in Standard Deviation Units of Raw Score, For Sponsor D Behaviorally Structured Approach



33

models have two sites which appear to perform better than would be expected without treatment on MAT Total Reading. Model C has one site and Model D two sites that are performing less well than would be expected in Reading. In Math, Models C and D also have one and two sites, respectively, performing less well than expected. With the exception of one site in Model D, all of the sites included in the behaviorally structured analyses perform less well in Spelling than would be expected without treatment.

In the cognitive-conceptual domain, the Raven's Coloured Progressive Matrices, a test of problem-solving ability, only one site in Model C performs better than expected. Otherwise, three and five sites, respectively, perform less well than would be expected without the program.

In the affective domain, the outcomes seen in Models C and D differ considerably. While both Models C and D show at least four sites where the Follow Through children perform better than would be expected on the Coopersmith test of self-esteem, only Model D shows any material and significant effects on the Intellectual Achievement Responsibility Scale (IARS), a measure of the extent to which children take responsibility for themselves or attribute their successes to the operation of internal or external forces. In the case of Model D, (Figure 12) site 006 performs better than expected on the IARS (+), and site 087 performs less well than would be expected. On the acceptance of responsibility for negative events, as noted by IARS₋ (-), in four sites Model D's Follow Through children performed better than would be expected. Only one site, 089, performs less well than would be expected.

In summary, both Models C and D appear to be achieving some of their intended goals, at least in Reading and Math. (Model C's Language goal is not included in this analysis.) However, Model D's Spelling goal does not appear to be met. While Model D's goal for increased social performance and behavioral growth is not assessed directly, there are indications that the children are performing better than would be expected in these areas.

When we change the criteria to the legal definition of educational disadvantage, i.e., performance which is one or more years below grade level, an additional perspective is added to our analysis. Conversion to MAT national norms shown in Figure 11 reveals that all sites in Model C are performing above the legal criterion of educational disadvantage in math, reading, and spelling at the end of third grade. For Model D (Figure 12) we see that, with one minor exception (site 089 math norm), all sites are also performing above the legal criterion. Since entry level is not controlled, it is problematic whether third grade MAT norm performance is attributable to treatment. However, while the pattern of effects for Model C appears to correlate with the socio-economic disadvantage of the site (as Figure 11 shows), this is not the case with Sponsor D (Figure 12). Here, most of the Follow Through sites are in the lowest income category (less than \$4,700) but the grade equivalents of the children are still at or near grade level.

To summarize, it appears that both behaviorally structured approaches are successful in meeting their goals. This is demonstrated by the fact that Follow Through children in many sites perform better than expected without the program in the basic skills of Reading and Math, and in the socio-emotional domain measured by the Coopersmith. In addition the FT children in Model D perform better than Model C's children on the IARS (-) in particular. This seems to indicate that the approach taken by Model D has generated within the children the recognition that they can take responsibility for both positive and negative events. Model C's and D's programs appear to differ primarily in the degree of individualization of instruction: They also differ in the type of children they serve, with Model D children being considerably more disadvantaged. Since both models appear to meet most of their goals, it appears that the behaviorally structured approach is generally successful with disadvantaged children. Beyond that, however, a high degree of individualization within a structured model may be particularly appropriate for severely disadvantaged children. Hence we conclude from the outcomes measured that the behaviorally structured approaches are generally appropriate models for compensatory education, and that while a group instructional approach may be appropriate for a moderately disadvantaged group, individualization of instruction may be necessary for a severely disadvantaged group.

Figure 11
 Grade Equivalents and Standardized Test Scores on MAT Subtests
 (Reading, Math, and Spelling) with Income Levels for
 COHORT II Sites in Sponsor C
 Behaviorally-Structured Approach

LEGISLATED CRITERION
 FOR DISADVANTAGE

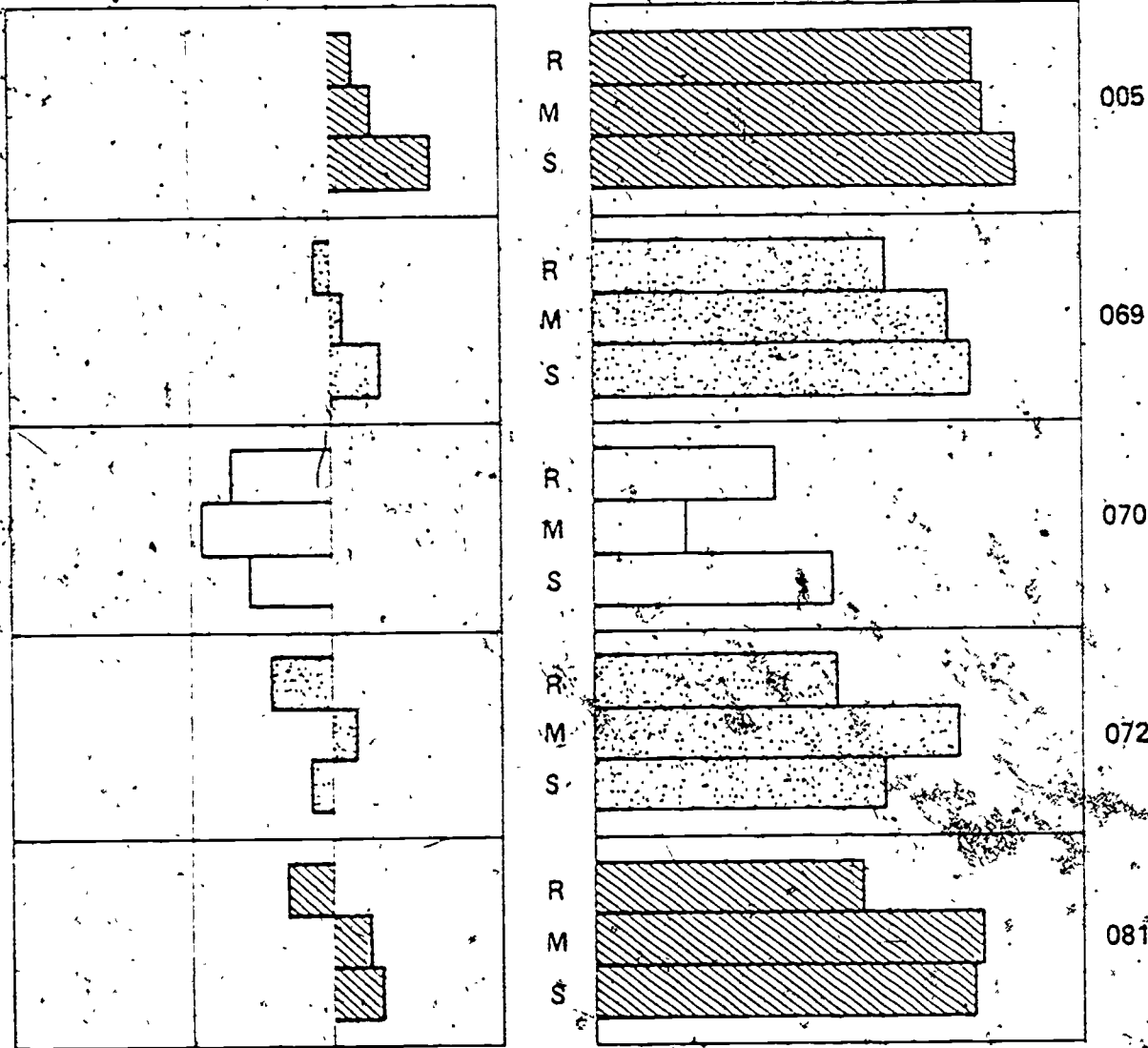
NATIONAL NORM LEVEL

GRADE EQUIVALENT

PERCENTILE OF MEDIAN

-2.0 -1.5 -1.0 -5 0 5 1.0

0 10 20 30 40 50 60 70



FT sites whose median income level exceeds \$6,000

FT sites whose median income falls between \$4,000 and \$6,000

FT sites whose median income is less than \$4,000

R = Reading, M = Math, S = Spelling

Figure 12
 Grade Level Equivalents and Standardized Scores on MAT Subtests
 (Reading, Math, and Spelling) with Income Levels for
 COHORT II Sites in Sponsor D
 Behaviorally Structures Approach

LEGISLATED CRITERION
 FOR DISADVANTAGE

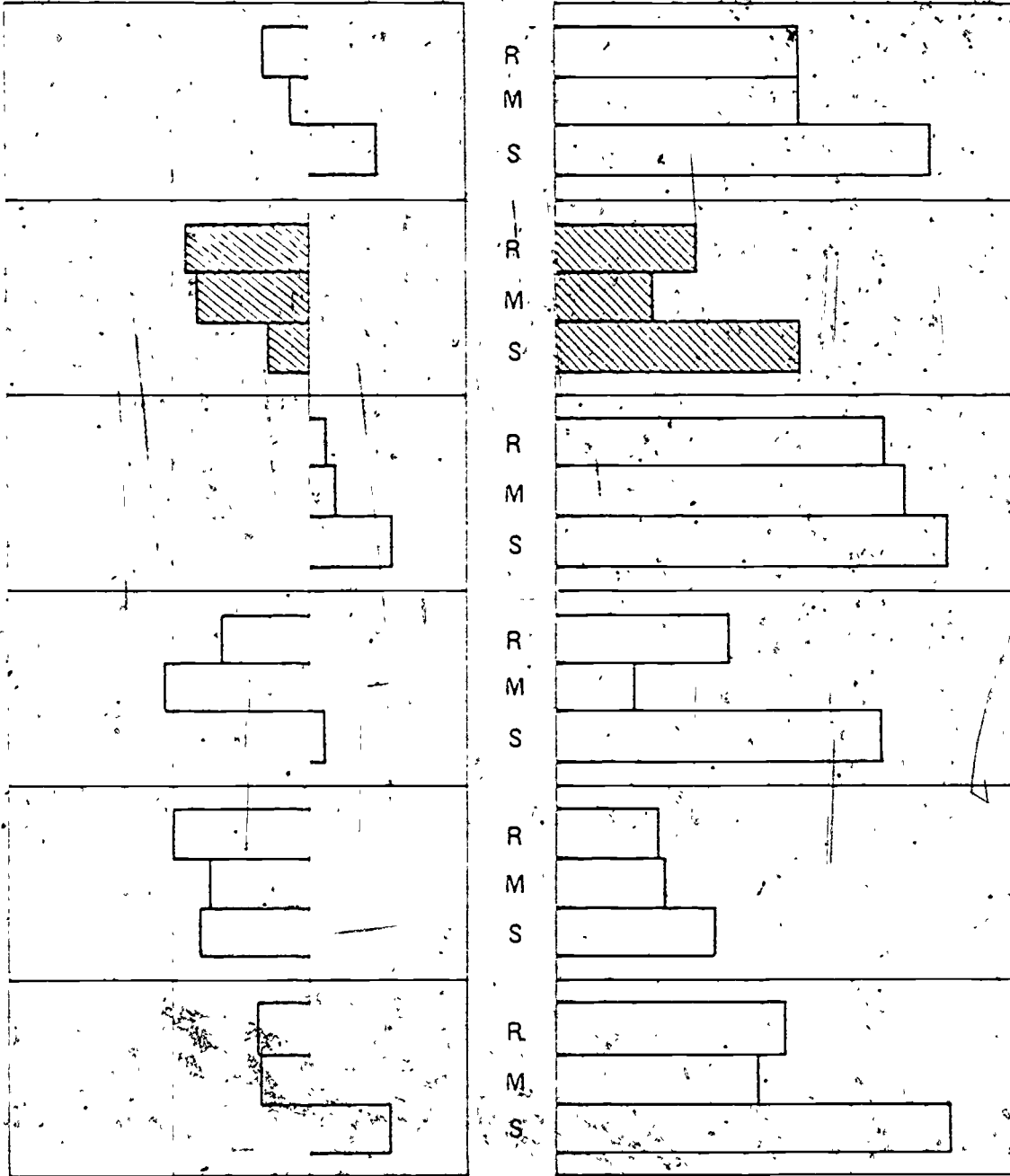
NATIONAL NORM LEVEL

GRADE EQUIVALENT

PERCENTILE OF MEDIAN

-2.0 1.5 1.0 0.5 0 0.5 1.0

0 10 20 30 40 50 60 70



□ FT sites whose median income level exceeds \$6,000

▨ FT sites whose median income falls between \$4,700 and \$6,000

□ FT sites whose median income is less than \$4,700

R = Reading, M = Math, S = Spelling

OPEN CLASSROOM CURRICULAR MODELS IN PROJECT FOLLOW THROUGH

Nancy Ames

Abt Associates Inc.
55 Wheeler Street
Cambridge, Mass. 02138

Overview of the Open Classroom Models

The Open Classroom models share a psychodynamic theory of learning. Their primary emphasis is on the development of "healthy" individuals with positive self-images and sound interpersonal relationships. The learning environment is flexible and supportive of interpersonal freedom and self-determination.

The Open Classroom models differ from the Behavioral models in both philosophy and approach. While the Open models value academic skills, which they see as essential for the optimal development of the "whole" child, these skills are seen as secondary in importance to a "healthy" sense of self and a positive attitude toward learning. The latter socio-emotional goals are seen as prerequisites for the former. Furthermore, the Open Classroom models favor an unstructured environment in which both teacher and child have a great deal of personal freedom, and the curriculum follows the children's interests and explorations. This contrasts sharply with the Behaviorist models in which teacher and student behaviors are predetermined, specific objectives are delineated, and instruction is carefully sequenced.

It is more difficult to distinguish between the Open Classroom and the Cognitive Development models. The philosophical difference between these approaches is not fundamental and several Follow Through models could be classified in either approach. Both approaches value the development of problem-solving skills as well as socio-emotional competencies. For the Open Classroom models, however, affective development is perceived as the primary objective. For the Cognitive Development models, the reverse is true.

For the purposes of this symposium, we have categorized two of the Follow Through models as exemplifying the Open Classroom curricular approach. We have labeled them E and F. Although both exemplify the Open

approach to education, there are differences between the models in theoretical origin, approach to educational innovation and change, content area emphasis, and curriculum offerings. There are also differences in each model's approach to structuring the learning environment.

Model E is based on 50 years of experience in providing education programs for young children. This model, which is designed to develop constructive, confident, productive, and coping human beings, attempts to develop the "whole" child. A teacher training program primarily, the model seeks to reach its objectives through improving the quality of teaching. Teachers are trained to use diagnostic tools to analyze child behavior, arrange the social milieu of the classroom, and provide instruction in accordance with individual learning needs. Teachers are given insights into how to enhance the capacity to probe, reason, and express feelings. A wide variety of materials is available for this model, including language arts and social studies programs. The social studies curriculum is designed to enhance interpersonal skills and develop cultural awareness.

Model F has its roots in the British Infant School approach to education. The primary objective of this model is creating a flexible, open, supportive environment in which the child's ability to learn is fostered by the educational setting. Since the model does not identify a single way of teaching children, a unique set of objectives, or a common set of materials, teaching behaviors, classroom environments, and materials vary widely. Interest areas are an important aspect of the learning environment in which children are free to explore, to choose, and to manipulate.

Sample

The sample children for all analyses reported here are Cohort II children, who have been in the Follow Through program from the beginning of kindergarten through the end of third grade. Six sites are included in the analysis for Model E. Four of these sites are large or medium-sized cities in the Northeast, one is a small, Northeastern community, and one is a medium-sized Southern community. Across all sites, children in this model are slightly below the Follow Through average on kindergarten entry scores, although they come from families with incomes slightly above the FT average. They do not differ significantly from the Follow Through group as a whole in terms of preschool experience or ethnic status.

There is great variability among Model E's sites however. One site is well above average on family income and entering kindergarten scores. Mothers in this site are better educated than the average for mothers of Follow Through children, and there are fewer female heads of households. One site is well below average on both family income and entering achievement. The four other sites are low on pretest scores but not on family income. Ethnically, the sites vary from 100 percent White to 100 percent Black.

Included in the analysis of Model F are three sites. Of these, two are large cities and the other a medium-sized city. All are in the Northeast or Northcentral region of the United States. Across all sites, children participating in this model were well below the Follow Through average in achievement upon entry to kindergarten, despite the fact that 91 percent had attended preschool. Family income across sites is slightly above the FT average. One hundred percent of the children are Black.

Model F's sites vary somewhat on both entry achievement status and family income. Two sites are close to the Follow Through average on pretest scores, one is far below average. On median family income, the former sites are above average, the latter below.

Teacher Reports as Indicators of Program Implementation

The data in Figures 13 and 14 illustrate the extent to which the teachers in each of the sites for these models characterized their classrooms as:

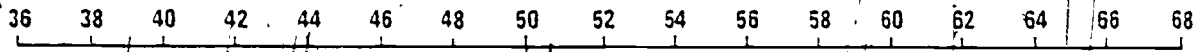
1. Teacher centered versus child centered;
2. Flexibly structured versus inflexibly structured;
3. Individualized versus group-oriented; and
4. Academically versus socio-emotionally oriented.

They also display teacher's satisfaction with the model.

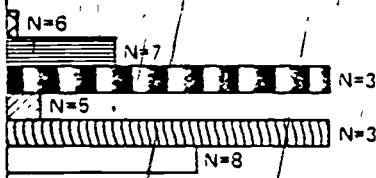
Model E's sites can be described as child-centered and flexible in nature -- although there is some variability among sites as to degree. Five of the six sites in this model are above average on individualization of instruction, while one is below. Sites are mixed in terms of their academic versus emotional emphasis. Teachers in two sites favor academic objectives; teachers in another two favor emotional development; and teachers in the other two show little preference. Finally,

Figure 13
Indications of Implementation From Teacher Questionnaire Clusters
for Sponsor E
Psychodynamic Approach

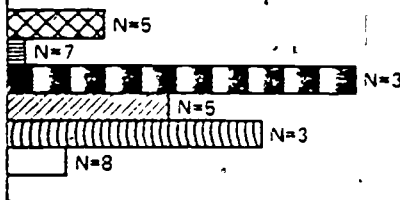
STANDARD SCORES SCALE
 Mean: 50 Standard Deviation: 10



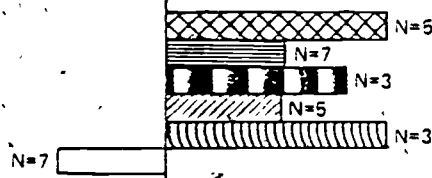
A. Teacher Centered vs. Child Centered



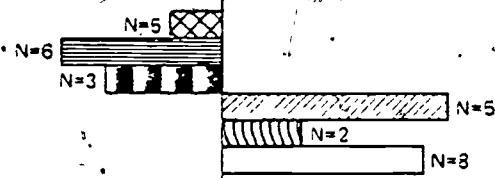
II. Flexible Classroom



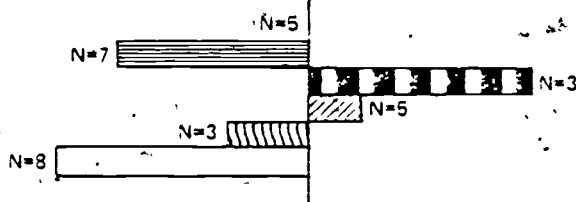
III. Individualization



IV. Academic vs. Socio-Emotional



V. Satisfaction with Sponsor

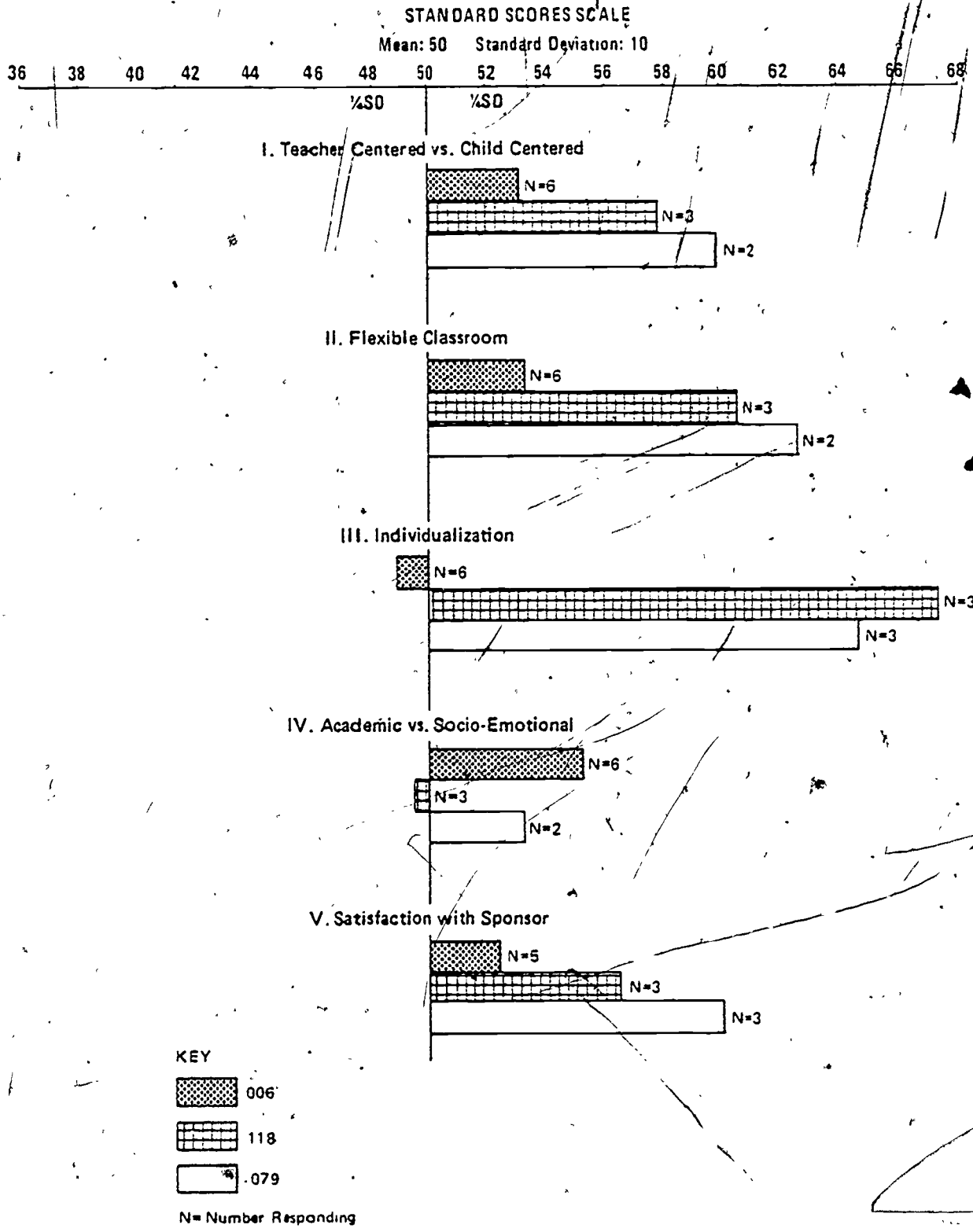


Key

- 005
- 006
- 050
- 052
- 054
- 056

N=Number Responding

Figure 14
 Indications of Implementation From the Teacher Questionnaire Clusters
 for Sponsor F
 Psychodynamic Approach



42

satisfaction with the model varies widely with one site extremely satisfied relative to teachers in other sites, two relatively dissatisfied, and the others neither satisfied nor dissatisfied.

Model F has two sites which are more child centered, flexible and individualized than most other Follow Through sites, according to teacher reports. The third is also child centered and flexible, but reports an average degree of individualization. Two of the three sites emphasize emotional rather than academic development; the third shows no clear preference. All three sites report above average satisfaction with the sponsor.

Overall, then, one could say that teacher reports reflect the 'models' educational philosophy and approach. Most variability is seen in the area of goals, with some teachers in both models reporting that their primary emphasis was on the development of emotional growth and others reporting that their primary focus was on the promotion of academic competencies.

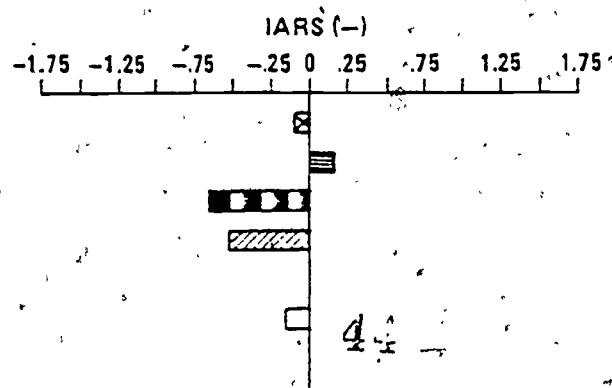
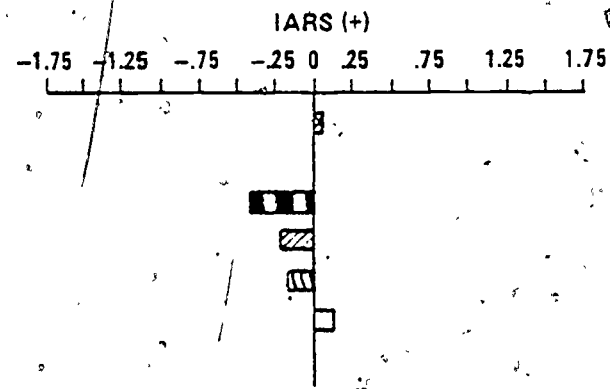
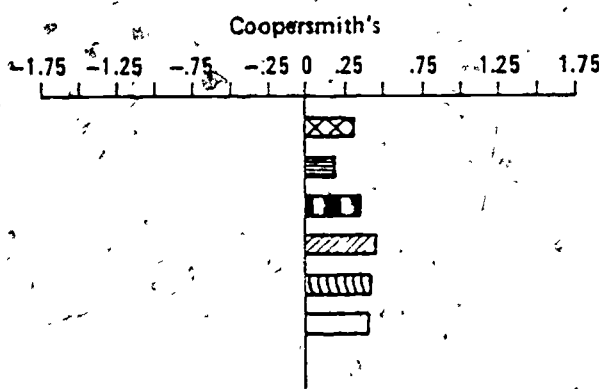
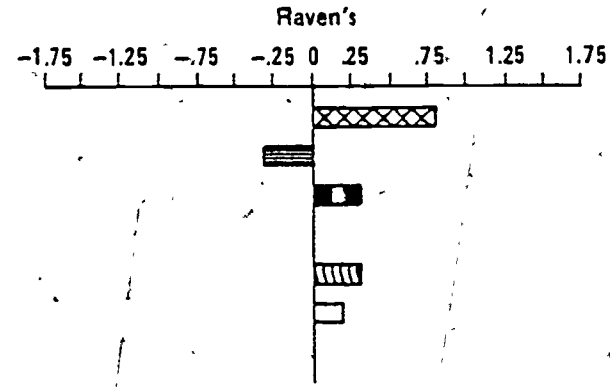
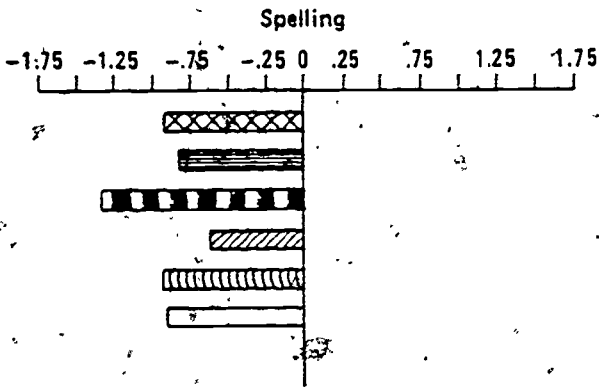
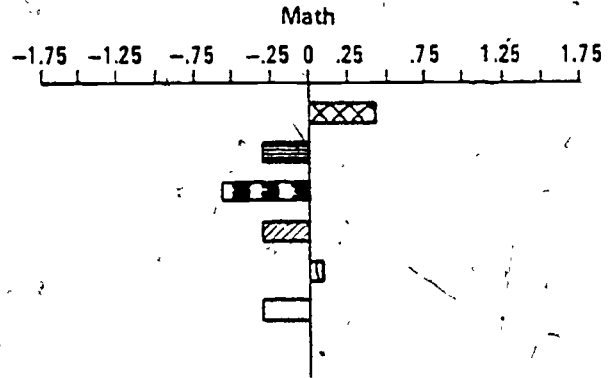
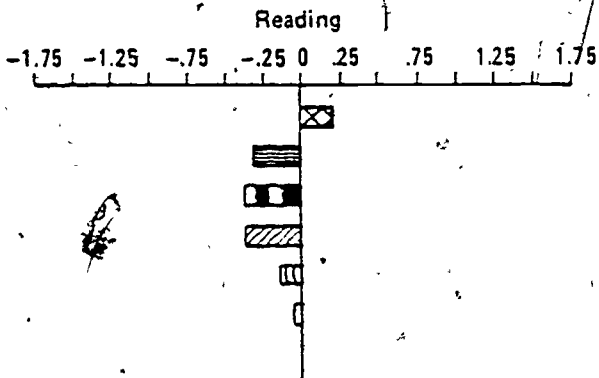
This inconsistency in goal statements may reflect varying degrees of implementation of the models, with some sites embracing the socio-emotional goals of the psychodynamic approach, and others preferring the traditional academic goals of the primary school program. Or it might reflect the philosophy of the approach itself. In focusing on the "whole" child, developing basic skills, problem-solving abilities, and personal-social skills are seen as intertwined.

Interpretation of Curricular Impact on Child Outcomes

Considering next child outcomes at the end of third grade, we examine curricular effects from two sets of data: first the contrasts of the Follow Through groups with the pooled non-Follow Through group, and second the contrasts of the FT groups with the national achievement norms.

As can be seen in Figure 15, Model E's effects vary by outcome and by site. On the Raven's Progressive Matrices test, two of the six sites performed better than would have been expected without treatment and only one performed below the expected level. In contrast, all of the sites scored below the expected level on the MAT Spelling outcome measure. Effects in Reading and Math varied, with sites performing above, below, and near expected levels. Finally, positive effects were found in one affective

Figure 15
 Sponsor Effects on Seven Third Grade
 Tests, Expressed in Standard Deviation
 Units of Raw Score, For Sponsor E
 Psychodynamic Approach



- 005
- 006
- 050
- 052
- 054
- 056

area. Four of the sites had a positive impact on self-esteem as measured by the Coopersmith Self-Esteem Inventory, and the other two sites scored at expected levels. On the internalization of responsibility for success measure, however, children in one site performed lower than would have been expected without treatment, and the others at expected levels. On the internalization of responsibility for failure measure, two sites performed below expected levels.

There is no clear pattern visible relating teacher reports and child outcomes for this model.

Turning to Figure 16, one can see that children participating in Model F performed at or below what would have been expected without treatment on all outcome measures.

To summarize the comparison of the Open or Psychodynamic curricular models with the pooled non-Follow Through group, we can make the following generalizations:

- In the cognitive and affective domains, these models do not produce similar effects. One model produces fairly consistent positive effects on the Raven's Progressive Matrices Test and the Coopersmith Self-Esteem Inventory. The other consistently does not.
- As for the academic area, there is variability between models and among sites. One model generally performs lower than expected in this domain and the other performs above expected levels in a limited number of sites.

The results of the regression analysis provide estimates of a model's performance compared to what would be expected without treatment. In contrast, grade equivalent scores provide information on the status of the Follow Through sites relative to the national norm group at the end of the children's tenure in the Follow Through program. Of course, the national norm group is a far less disadvantaged group. Follow Through students generally would be expected to perform well below this group without treatment. To the extent that FT children are at or near grade level at the end of third grade, Follow Through might be said to be having some impact. Also, if FT appears to interrupt the strong relationship between income and achievement, then FT may be said to be having an effect. As can be seen in Figure 17, Model E has two sites one-half year or less below grade level, two sites less than one year below grade level, and two more sites more than one year below grade level. Thus in four of the

Figure 16
 Sponsor Effects on Seven Third Grade
 Tests, Expressed in Standard Deviation
 Units of Raw Score, For Sponsor F
 Psychodynamic Approach

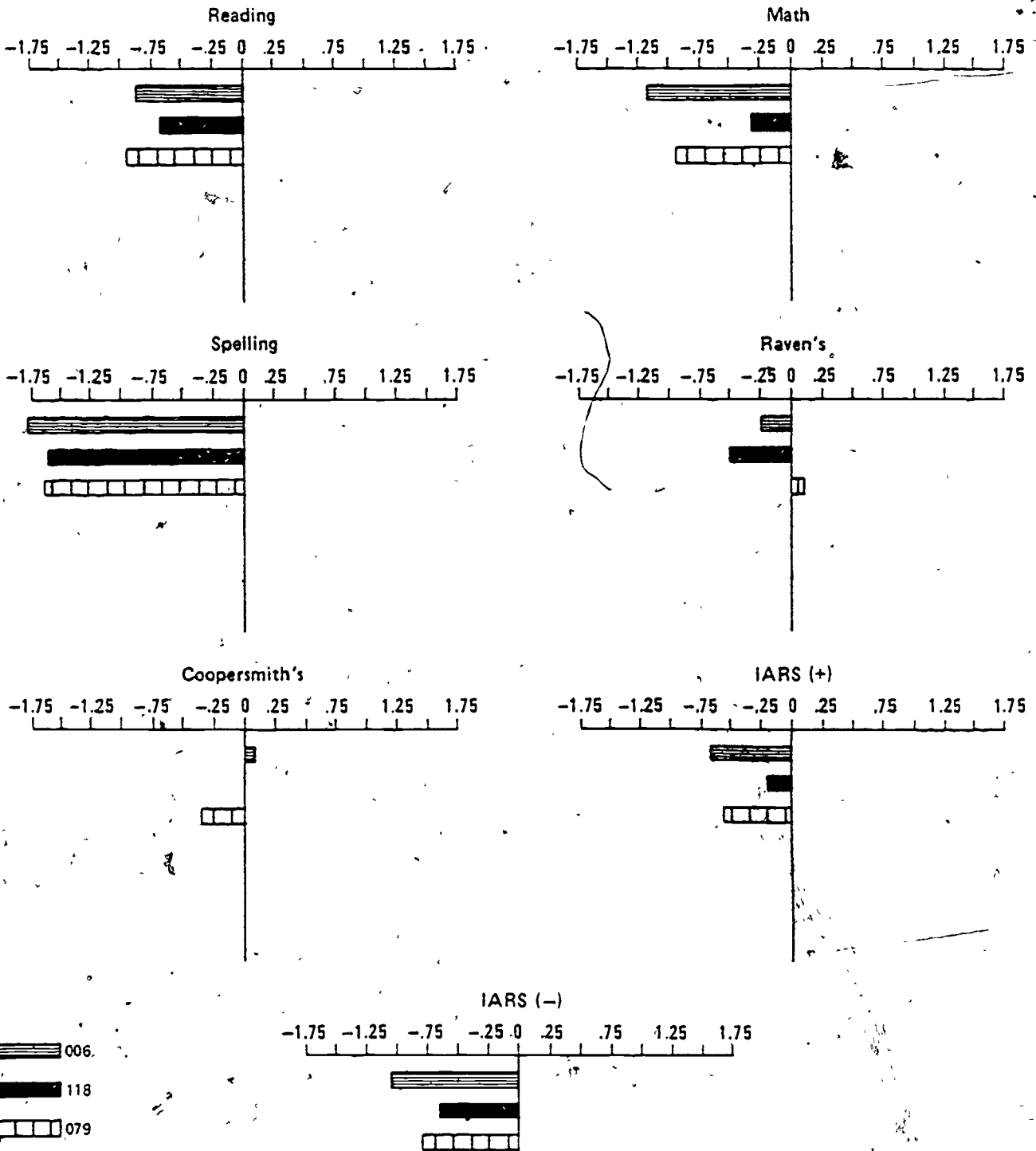
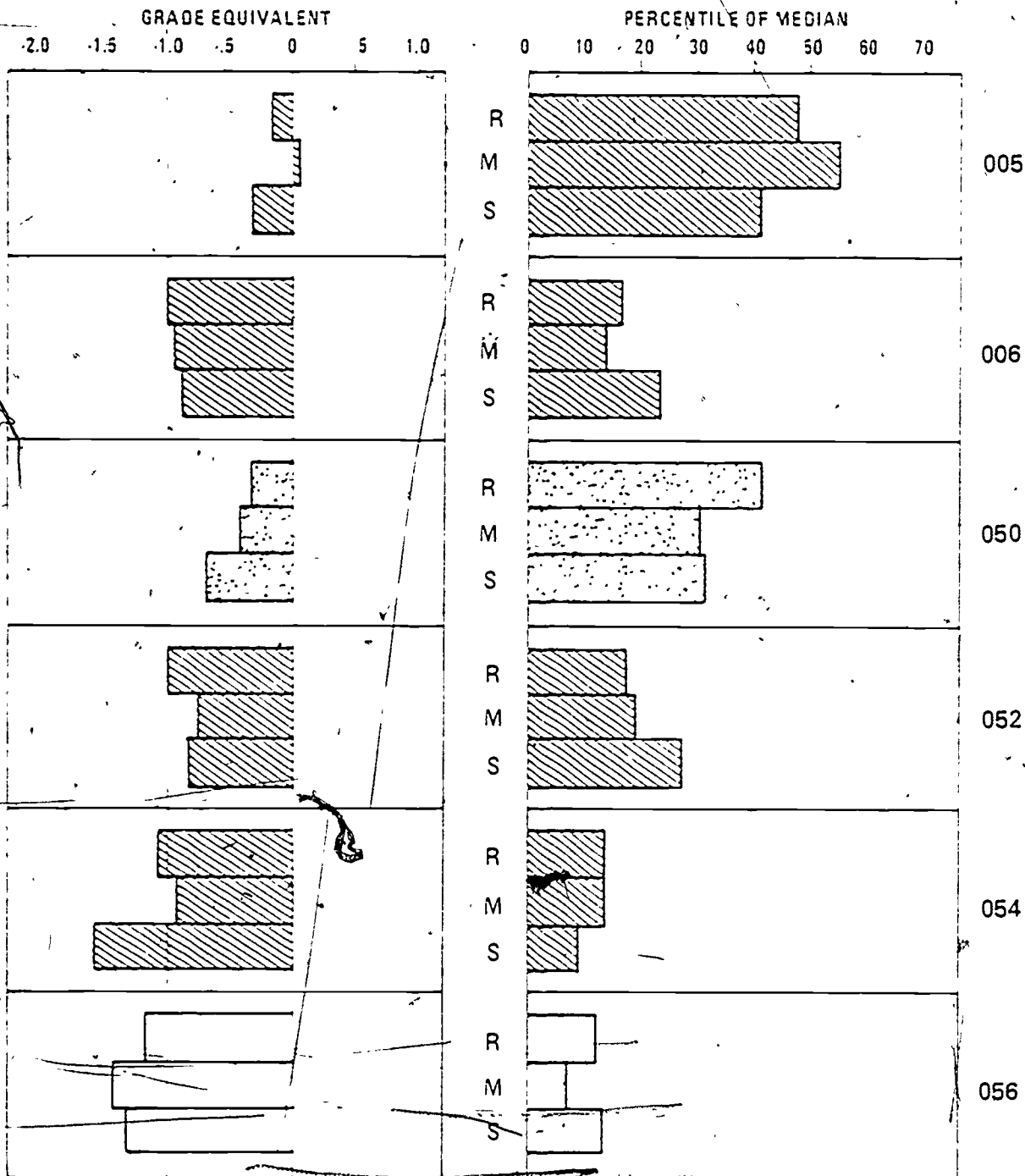


Figure 17
 Grade Equivalents and Standardized Test Scores on MAT Subtests
 (Reading, Math, and Spelling) with Income Levels for
 COHORT II Sites in Sponsor E
 Psychodynamic Approach

LEGISLATED CRITERION
 FOR DISADVANTAGE

- NATIONAL NORM LEVEL



R = Reading M = Math S = Spelling

six sites children in this model would not be classified as legally disadvantaged. Since one of these sites is also relatively poor in median income, Follow Through may be said to be having a particularly strong impact in this site.

In contrast, Figure 18 reveals that Model F has two of its three sites performing one or more years below grade level. The other site is more than one year below grade level on one of the outcome measures, less than one year below on the other two. Figure 18 also shows that there is a correspondence between grade level performance and site mean income. Thus, there is no evidence in these data that Model F has disrupted the poverty/achievement cycle in these sites.

Qualitative Assessment of Curricular Impact

Comparisons with either a pooled non-Follow Through group or a national norm group are by nature sensitive to the differences between the Follow Through and comparison groups and the nature of the FT group itself. These and other factors often make it difficult to separate treatment effects from experimental error. The power of our tests is such that effects must be quite large to be observed. If we ignore for a moment the size of the effects, and examine only the pattern or ordering of effects across domains, we find a somewhat striking pattern (see Table 1).

The Open approach appears to have its greatest effects in the affective domain, particularly in the area of self-esteem. It has its next strongest effects in the problem-solving area. Its weakest effects are in Spelling, a somewhat rote learning task.

As pointed out in the introduction of this paper, it is difficult to separate the Open or Psychodynamic approach from the Cognitive Discovery approach. Both value an open, manipulative environment in which children are encouraged to learn how to learn. This philosophical closeness is clearly reflected in the results for this approach -- the strongest treatment effects appear to be in the cognitive and affective areas.

Figure 18
 Grade Equivalents and Standardized Test Scores on MAT Subtests
 (Reading, Math, and Spelling) with Income Levels for
 COHORT II Sites in Sponsor F
 Psychodynamic Approach

