DOCUMENT RESUME

ED 332 389 EC 222 019

AUTHOR	Vye, Nancy J.; And Others
TITLE	Dynamic Assessment of Intellectually Handicapped Children. Alternative Assessments of Handicapped Children. Technical Report No. 4.
INSTITUTION	George Peabody Coll. for Teachers, Nashville Tenn. John F. Kennedy Center for Research on Education and Human Development.
SPONS AGENCY	Department of Education, Washington, DC.; National Inst. of Child Health and Human Development (NIH), Bethesda, Md.
PUB DATE	Oct 85
CONTRACT	G0083C0052; HD15052
NOTE	58p.; Print is small and may not reproduce well.
PUB TYPE	Reports - Research/Technical (143)
EDRS PRICE	MF01/PC03 Plus Postage.
DESCRIPTORS	Ability; Elementary Secondary Education; *Evaluation Methods; *Mediation Theory; *Mental Retardation; Predictive Measurement; *Prompting; Research Needs; Research Projects; *Student Evaluation
IDENTIFIERS	*Dynamic Assessment

ABSTRACT

This paper reviews research literature and research projects of the John F. Kennedy Center for Research on Education and Human Development, on the topics of the utility of dynamic assessment for predicting learning ability, its utility for generating educationally relevant prescriptions, and factors associated with the implementation of dynamic procedures. Following an overview of research issues, a "continuum of assessment services" model is presented. Along the continuum are such forms of dynamic assessment as graduated prompting and mediation assessment. In a discussion of the utility of dynamic assessment for predicting learning ability, static and dynamic assessment are compared, the relation between dynamic assessment and within-and-across-domain transfer is discussed, and directions for future research are outlined. A section on dynamic assessment and educational prescriptions contains observations of young children's learning and directions for future research. Factors related to the implementation of dynamic assessment are then explored. It is concluded that static and dynamic measures produce different estimates of learning, that dynamic assessment is useful for predicting learning, that mediation assessment is effective for remediating strategies that help children improve their performance, and that dynamic assessments can have important effects on teacher expectations. (JDD)

ALTERNATIVE ASSESSMENTS OF HANDICAPPED CHILDREN

A Series of Technical Reports and Working Papers

Technical Report No. 4

Dynamic Assessment of Intellectually Handicapped Children

Nancy J. Vye, M. Susan Burns, Victor R. Delclos, and John D. Bransford

John F. Kennedy Center for Research on Education and Human Development
Peabody College, Vanderbilt University

ALTERNATIVE ASSESSMENTS OF HANDICAPPED CHILDREN is a series of Technical Reports and Working Papers designed as a vehicle for timely dissemination of work-in-progress dealing with the assessment of handicapped children. Primary support for this series is provided by Grant No. G0083C0052 awarded to Vanderbilt University by the U.S Department of Education. Additional support for production of the series is provided by Grant No. HD15052 awarded to Vanderbilt University by the National Institute of Child Health and Human Development.

Technical Report No. 4 (October, 1985)

Dynamic Assessment of Intellectually Handicapped Children[1]

Nancy J. Vye

University of Western Ontario

M. Susan Burns

Vanderbilt University

Victor R. Delclos

Vanderbilt University

John D. Bransford

Vanderbilt University

Running head: Dynamic

To appear in C. S. Lidz (Ed.), Dynamic assessment: Foundations and fundamentals. New York: Guilford Press, in press.

The purpose of our chapter is to present some preliminary findings from our research on dynamic assessment. The research project has been in existence for a year and half, and while the data do not address many points, they nevertheless provide an excellent context within which to discuss issues related to learning assessment.

Our research project, and hence our discussion, is organized around three topics: the utility of dynamic assessment for predicting learning ability, its utility for generating educationally relevant prescriptions, and factors associated with the implementation of dynamic procedures.

Our thinking on dynamic assessment has been influenced in fundamental ways by the ideas of Feuerstein and his colleagues (Feuerstein, Rand, & Hoffman, 1979), Brown, Campione and their colleagues (Brown & French, 1979; Campione, Brown, & Ferrara, 1982), and Vygotsky (1962, 1978). It is beyond the scope of this paper to present the theoretical ideas of these authors in any detail. Instead, we refer the reader to relevant papers in this volume (see Feuerstein, Brown and Campione and Minick chapters). Let us say at the outset however that we are fully responsible for any departures from and/or misconstruals of this theory base.

1.0 Overview of Research Issues

When we first began thinking about how to give shape to a research project on dynamic assessment, it seemed important that our research address questions in the three broad areas of identification, prescription and implementation. The identification area seemed important for several reasons: The first related to our initial focus on preschool children. It is well-documented that until children are about 5 years of age, static assessments of learning ability are relatively unreliable (Brown & Ferrara, 1980; Lidz, 1983; Reynolds & Clark, 1983; Simner, 1983). Thus, it seemed to us that dynamic assessment measures would meet an important need if, as suggested by our initial data and the data of others (Brown & Ferrara, 1980; Budoff, 1967), they proved to be a valid means of identifying young children with learning difficulties.

A second reason for focusing on the identification issue was to establish the relation between estimates of learning ability derived from static assessments and those derived from the dynamic assessments we had developed, and perhaps more importantly, to determine the validity of these dynamic estimates. Consider, for example, a situation in which a static measure of learning ability fails to predict children's performance in dynamic assessment. To what might this be attributable? One plausible explanation is that the inconsistency is due to the unreliability of one or both of the measures. We, of course, want to rule out such explanations. It seemed important therefore to begin investigating the extent to which dynamic assessment estimates were predictive of performance on tasks in related and different domains.

A second area that we wished to address in our research was prescription. While most would agree that it is important to go beyond mere classification, standardized diagnostic/prescriptive procedures are quite rare and the outcome data on those available have been disappointing (Arter & Jenkins, 1979). While our work in this area is just beginning, our approach appears promising for several reasons: In a dynamic assessment, one is able to gather information about the effectiveness of various instructional manipulations. Furthermore, dynamic assessment emphasizes learning processes rather than underlying "abilities" or "traits". We assume that task-relevant knowledge and general and task-specific strategies can be assessed in dynamic assessment, are amenable to change, and that these changes will produce meaningful improvements in task performance. This information in turn might prove useful to teachers.

In the third area of implementation we have been addressing the question of whether there is an association between dynamic assessment and a change in perception about a child's abilities and potential. We have also been investigating whether teachers see the utility of information provided by a dynamic assessment, and finally, ways to communicate dynamic assessment prescriptions to facilitate the likelihood of their implementation.

The theoretical rationale for our project revolves around a "continuum of assessment services" model. The continuum model involves initial screening by means of an individually administered intelligence test. Children whose IQ results are in or above the average range would not receive any further assessment, while those who scored more than 1 SD below the mean would receive "graduated prompting" dynamic assessment in each of the verbal, quantitative and perceptual performance domains. (Graduated prompting is described below.) Children who are above criterion on these tasks would be viewed as responsive to instruction, and no further assessment would be provided. A young child who does poorly on static measures, but performs at high levels following graduated prompting may well be a child whose primary need is an enriched learning environment such as that provided by a preschool experience. A school age child who shows this profile may be one who could profit from the benefits of a resource person in addition to regular classroom instruction. Our assumption here is that these children are able to learn, but lack experience with the concepts and procedures or the motivation required to do well on intelligence tests.

Graduated prompting was placed second along the continuum because its instruction seems most similar to the type one might see in a classroom. If this assumption about graduated prompting is correct, then we would be justified in using it to predict the children who could profit from regular or enriched classroom experiences. From our perspective at least, the assumption has face validity. For example, graduated prompting instruction is only as explicit as needed, and classroom group instruction, by necessity, is often just sufficient to enable independent performance in most learners.

There are several additional reasons favoring the use of graduated prompting at this stage: It is a fully-scripted procedure, and hence may not need to be administered by a professional. This may make it more cost effective and easier to implement on a broad scale. Furthermore, one can derive measures of learning speed (i.e., number of prompts to criterion) from graduated prompting, and learning speed may prove to be a particularly sensitive predictive measure.

Children who are below criterion following graduated prompting would next receive "mediation" dynamic assessment. (This procedure is also discussed more fully below.) Children who reach this final stage of assessment would presumably be those experiencing fundamental learning difficulties. The purpose of this assessment would be to determine the child's ability to profit from a period of intensive, contingent instruction, that is, instruction whose content is determined by the needs of the individual child. Another purpose would be to derive educationally relevant descriptions of learning processes. These descriptions -- which we will discuss more fully below -- would entail information about content and processes on which the child was experiencing difficulty, as well as effective remediation strategies.

2.0 Mediation Assessment

Our continuum of assessment services model employs assessment procedures based on Feuerstein's mediation principles, (i.e., mediation assessment) and

Brown and Campione's zone of proximal development procedure (i.e., graduated prompting assessment). In order to make our research feasible, we were compelled to develop a brief form of Feuerstein's dynamic assessment. His extended assessments can last from a number of hours to several days, while our mediation assessments last only about 30 minutes. Despite the brief nature of the assessment, our results confirm that a session of this duration can provide valuable information about children's learning. Where more indepth information is needed, for example about specific prescriptions and their implementation, we have used a series of mediation sessions (i.e., as in the case of the single subject research described later).

Mediation procedures were developed initially for children of preschool age in the perceptual performance (PP) domain (Burns, 1985), and the task used was an adaptation of the Representational Stencil Design Task (Arthur, 1947; Burns, Haywood, Delclos, & Siewart, 1985). Since then, we have designed tasks and procedures for young children for the quantitative (hereafter called quantitative task or QT) and verbal domains, and for middle school age children in the core curriculum areas of mathematics (for both computation and word problems) and reading comprehension.

Consider for a moment the mediation procedure for the Stencil Design Task (SDT). On the SDT, the child is presented with an array of 18 colored cards. Twelve of the cards have a shape or pattern cut out of them ("cut-outs"); the remaining 6 cards are of a solid color ("solids"). The child's task is to place a cut-out on top of a solid to create a design that matches a model design. During mediation, the child receives training on 4 such designs. For research purposes, children are posttested on 8 new designs which they are asked to complete without assistance from the examiner.

The components of the mediation assessment for the SDT are representative of those used in other task domains. Appendix A describes the structure of the assessment. Three general instructional components are included: familiarization of materials and basic cognitive functions, instruction on task-specific rules and procedures, and feedback. As illustrated in Appendix A, familiarization of materials for the SDT involves helping the child distinguish the relevant dimensions of color, shape, size, location and orientation, while familarization of basic cognitive functions involves encouraging the child to compare cards on the basis of these dimensions and to systematically search the entire array of cards. Feuerstein (Feuerstein et al., 1979) has suggested that children who have been deprived of sufficient mediated learning experiences are often deficient with respect to these fundamental learning processes or what he calls "basic cognitive functions", and this in turn interferes with their ability to perform more complex cognitive operations. Thus, an important aspect of the mediation assessment is to encourage the child to engage in these behaviors, and for the examiner to communicate this expectation prior to specific instruction on the task.

Task specific rule instruction involves teaching the child the need for two cards, the order for combining cards, the effect of combining different cards, and the use of the model. As we shall see, the graduated prompting procedure consists of instruction on these same rules, although the method of instruction differs.

The third component of the mediation assessment is feedback. After the child has completed a stencil design, the examiner draws the child's attention to and discusses both the errors and the correct aspects of the construction. The feedback portion of the assessment is perhaps the most informative with respect to diagnosing strengths and weaknesses in the child's approach and the child's response to different attempts at remediation.

3.0 Graduated Prompting Assessment

The procedures used in the graduated prompting assessment are based directly on the work of Brown, Campione and associates (Brown & Ferrara, 1980; Bryant, 1982; Campione, Brown, Ferrara, Jones, & Steinberg, 1983). In graduated prompting, a series of prompts or hints is used to teach the rules needed for task completion. The prompts are ordered in explicitness, with general prompts given first and more explicit prompts later. One prompt is presented to the child each time s/he cannot complete the task.

The sequence of prompts developed for the SDT are contained in Appendix B. Notice that the initial prompts are very general; the child is first reminded about his/her previous construction. Subsequent prompts teach the 2-card rule, then how to identify the solid and cut-out in the model, how to combine cards, how to search for cards in the array, and so forth. In our research using the SDT, graduated prompting is given on 4 designs and unassisted, posttest performance is assessed on 8 designs. Thus far we have developed graduated prompting procedures for young children in the perceptual performance and quantitative task domains.

4.0 The Utility of Dynamic Assessment for Predicting Learning Ability

4.1 Comparison of Static and Dynamic Assessment

When we were initiating our work, the available research on dynamic assessment suggested that dynamic measures were capable of further discriminating differences in learning ability among children who had been homogeneously grouped on the basis of static intelligence scores (Brown & Ferrara, 1980; Budoff, 1967). The correlations that had been reported between various intelligence measures and learning measures (i.e., performance gains, learning speed, transfer) were either low or non-significant. However, full scale IQ was typically the static measure employ... d it could be argued that the reported correlations were spuriously low because of the use of a global static measure (compounded in some cases by attenuation due to restriction of IQ range). We wanted therefore to establish the relation between dynamic measures and "non-global" (i.e., subscale) static measures. In addition, we wanted to extend the findings to our new tasks and dynamic procedures.

Thus far in our research we have not found evidence to suggest that the previously reported results are in any way spurious. In our first experiment, hereafter Study A, we calculated correlations between dynamic measures and full- scale and subscale scores. The McCarthy Scales of Children's Abilities (1972) were used as static measures. The dynamic measure was children's unassisted performance on the SDT following graduated prompting. Table 1 contains the obtained correlations. The pattern of results is quite clear:

Insert Table 1 about here

Consistent with previous reports, General Cognitive Index (GCI) bears only a moderate relationship to measures on the SDT following dynamic assessment. While the correlation did not achieve significance, we suspect that the obtained score is an underestimate that is most likely due to a restriction of range in GCI scores.

The correlations between the subscale and the dynamic measures were of particular interest in Study A and inspection of Table 1 reveals that the correlation is significant when the task domain is the same (e.g., SDT and PP). While this correlation is higher than the GCI/SDT correlation, it is clear that a child's subscale score would not be a very reliable indicator of his/her response to instruction.

A slightly different breakdown of Study A makes the point more forcefully. The 77 handicapped children--the number includes children from a second study, Study B, as well--were grouped according to GCI. Groups consisted of children whose GCIs were below 37, from 37 to 52, from 53 to 68 and, from 69 to 108. The handicapped children in the 69-108 are those children who have a 2 standard deviation difference between subtests - many are from a school for learning disabled children. Figure 1 represents the percentage of children in each GCI group who learned to do the SDT following dynamic assessment. As you can see, a substantial number of children reached criterion on our brief dynamic assessment procedure: 36% of the children with IQs between 37 and 52 reached criterion; 53% of the children with IQs between 53 and 68; 82% of the children with IQs between 69 and 108; and even in our lowest IQ group (below 37), 26% reached criterion. In spite of their static classification, a sizable number of children in each group were responsive to instruction.

The results of Study A and B point to several conclusions: Consistent with previous work, full scale intelligence measures do not predict dynamic performance with any great precision. More significantly, the present research indicates that subscale measures increase prediction only slightly. A large proportion of our children would have been misclassified with respect to their ability to profit from instruction if static measures had been used as the basis for classification. It is also noteworthy that these results were obtained for the first time using our adapted mediation assessment and thus are an initial demonstration of the feasibility of the procedure.

4.2 The Relation Between Dynamic Assessment and Within-and Across-Domain Transfer

Our long range purpose in conducting experiments such as Study A and Study B is to investigate the predictive utility of dynamic assessment procedures. We hope over the course of our work to demonstrate that dynamic assessment can serve as a valuable tool either alone or in combination with static, normative measures for estimating children's future learning on a task. A first step in the process is to establish that dynamic estimates are not redundant with static estimates (i.e., perfectly correlated). Study A and B accomplish this goal. Although strictly speaking, the low correlations that we reported could be attributed to the unreliability of either or both the static and dynamic measures, the consistency with which moderate correlations are observed argues against the unreliability hypothesis.

A second step in the process is to explore the relation between dynamic assessment performance and performance on transfer tasks within and across task domains. In other words, we need to assess the concurrent and predictive validity of dynamic measures. We have data from Study A and Study B and from two single-subject projects that begin to address these issues. While some of our data analyses are in preliminary stages, particularly in the case of Study B, there is sufficient information available to suggest that performance following dynamic assessment is predictive of within-domain transfer performance, but not predictive of across-domain performance.

Consider Study A once again. The design is one in which children first receive a session of dynamic assessment on the SDT, either mediation or graduated prompting, followed by unassisted or independent performance (hereafter IP) on the SDT and on three, within-domain transfer tasks. Table 2

Insert Table 2 about here

contains children's scores on each of the tasks and illustrates that there is a relation between performance following dynamic assessment and performance on the transfer tasks. Children who do well on the SDT, and by this we mean children who meet or exceed a criterion of 75% accuracy, also do well on two other stencil tasks (one involving stencils of animal shapes and the other involving the same geometric shapes as the SDT but in an appropriate figure-ground relationship), and the Animal House Coding Task (raw score). Similarly, children who do poorly on the SDT tend to do poorly on the transfer tasks.

The differences on transfer performance between above and below criterion children are statistically significant when considering the graduated prompt conditions. The transfer differences did not reach significance in the mediation groups, although the means are in the same direction. The mediation results are most likely due to the relatively good transfer performance of the below criterion children, and suggest that mediation instruction promotes greater generalization. Although we have not examined our ideas empirically we suspect that the mediation and graduated prompt procedures may differ in terms of the relative emphasis placed on metacognitive skills. The mediation procedure, for example, devotes instructional time to search and self-checking strategies. Transfer of strategies such as these may account for the better performance of the below criterion mediation children over their graduated prompting counterparts. Before leaving this issue -- we will return to it shortly -- we should say that the observed differences in transfer performance do not imply that one procedure is "better" than the other; the value of each procedure is established by reference to the function it is designed to serve. As suggested earlier, the graduated prompting procedure may prove to be more valid than the mediation procedure when used to predict a child's response to classroom instruction, assuming that classroom instruction is more similar to the graduated prompting than the mediation type of instruction.

The results of Study A suggest that dynamic assessment, that is, instruction of a brief duration, may prove helpful for predicting how well children will perform on related tasks. We have also obtained similar results with more severely handicapped children. Because of the degree of handicap, we have used a single-subject research methodology, and have extended the mediation over a number of sessions. In the first experiment, baseline measures consisting of IP on the SDT and the two stencil transfer tasks were collected over a series of sessions (the actual number was staggered across the 3 children who participated). The mediation sessions followed, and continued until such time as an improved and stable level of performance was observed on the SDT (IP measures on SDT were taken at the end of each mediation session). At this point, mediation instruction was discontinued, and maintenance of learning was assessed over several more sessions. Figure 2

Insert Figure 2 about here

depicts the SDT results for the 3 children (fictitious names are used to protect the children's identities). The results of particular interest here are the transfer results. Independent performance on these tasks was assessed in a session following each of the mediation sessions. Figures 3 and 4, which

Insert Figures 3 and 4 about here

depict SDT and transfer results of two of the children, show quite clearly the spontaneous transfer of SDT learning to the other stencil tasks. Again, this illustrates the correspondence between performance following dynamic assessment and performance on related tasks. These results are in contrast to much of the training research conducted with mentally handicapped persons wherein transfer of learning has been the exception rather than the rule. The present findings may be due in part to the multi-session nature of our research design. There does seem to be a time lag between improvement on the criterial task and improvement on the transfer task. These latter improvements would be missed if, as is usually the case, only a single transfer session was used.

Having spoken about within-domain performance, let us address briefly the case of across-domain performance. We have conducted two experiments, Study B and a second single-subject design, that bear on this issue, and in both cases we looked at the correspondence between performance on a PP and a Q task. In contrast to the results from within-domain tasks, we have not found evidence of spontaneous transfer. Nor have we found any evidence to suggest that responsiveness to instruction in one task domain predicts responsiveness to instruction in a second. That is, the children who achieve our learning criterion following dynamic assessment (either mediation or graduated prompting) on the SDT do not consistently achieve criterion following instruction on the QT. Nor are the below criterion SDT children always below criterion on the QT.

To summarize then, our work thus far suggests that there is a relation between performance following dynamic assessment and performance on tasks in a related domain. Study A showed that handicapped children who met a learning criterion tended to do well on a series of near transfer tasks, while those children who did not meet the criterion did not transfer nearly as well. In contrast, we have not found evidence indicating across domain correspondence. Study B failed to show that classifying handicapped and non-handicapped children according to learning status on the SDT improved appreciably our ability to predict learning status on the QT. We should make the disclaimer however that we have yet to complete correlation analyses on Study B data; what we have presented are the results of preliminary chi-squares. Nonetheless, we expect that this will leave our major conclusions unchanged.

The within domain transfer observed in the single subject experiment deserves further comment. At first glance it might seem as if the result contradicts our conclusion that performance following dynamic assessment helps us to predict performance on similar tasks since the children, who were selected because of the severity of their handicap, showed evidence of spontaneous transfer. The result also seems to fly in the face of other reports (Belmont & Butterfield, 1979; Brown, Bransford, Ferrara, & Campione, 1984; Campione et al., 1983) indicating the utility of learning and transfer profiles for identifying children who are "truly" mentally handicapped. These not withstanding, we suggest that our conclusion holds since the children did perform poorly at first both in response to mediation, and on the transfer tasks. It was only after several intensive mediation sessions that children began to improve on the SDT, and the transfer effects were delayed by several more sessions. Nonetheless -- and we wish to stress the point because we fear that it is often lost -- the children did possess the ability to learn the task if provided with the appropriate intervention, and were able to transfer the acquired skills. Conclusions, ours and those of others, about "non-learners" or "non-transferers" need to be interpreted cautiously, that is, in relation to the learning criteria and tasks selected by the experimenter and not in an absolute sense.

4.3 <u>The Utility of Dynamic Assessment for Predicting Learning Ability: Directions for Future Research</u>

Our research suggests several conclusions about the utility of dynamic measures: First, Study A and B results indicate that static and dynamic measures do indeed produce different estimates of learning, and that this discrepancy is not an artifact of taking a fullscale rather than a subscale measure. Furthermore, there is an ever-increasing body of evidence that argues against the possibility that this discrepancy is attributable to the unreliability of either the static or dynamic measures. The stability of the static/dynamic correlations across replications weakens any claim about unreliability.

A second conclusion that emerges from our research relates to the validity of dynamic measures. (It is important of course to establish that the information derived from our dynamic assessments, its uniqueness and reliability notwithstanding, will help us to predict learning and performance.) The results of Study A indicate that responsiveness to instruction on the SDT is predictive of performance on transfer tasks, a striking correspondence in view of the fact all of the children had been below criterion during the first phase of the study and a correspondence that bolsters validity claims. Our single subject work leads to a similar conclusion. The learning curves obtained from the mediation closely parallel (with a time lag) those obtained from untrained transfer measures, indicating that a child's response to instruction on one task is closely tied to uninstructed learning in another. Taken together the data represent evidence of the concurrent validity of dynamic assessment.

The above mentioned results were derived from tasks selected from the same domain, that is, the perceptual performance domain. The across-domain data suggest a different conclusion. We have not found evidence of between domain transfer (between Q and PP domain) in our single subject research. Nor do Study B data show a relationship between responsiveness to instruction in one domain and responsiveness in another. As far as we can tell then, dynamic measures, if and when impl-mented, would need to be taken in each domain for which learning is to be predicted.

One final point deserves reiteration and that is the relation between graduated prompting and mediation assessment. The data from Burns (1985), Study A, and Study B show that the two procedures do not produce different numbers of above- and below-criterion children, and as such, appear to be equally effective for teaching the task at hand. The mediation procedure does seem to be somewhat more effective for promoting transfer. We are doubtful however that this will prove to be an important factor in determining their relative predictive utility because what's important for prediction is a rank ordering of children's performance and not the absolute level of performance. As discussed earlier, we do believe that the two procedures may serve different assessment functions. The GP procedure by virtue of its scripted nature seems ideally suited for prediction. As suggested by the work of Brown and colleagues (Bryant, 1982; Campione et al., 1983), using a measure of the number of prompts needed to reach criterion in conjunction with static measures (instead of the pre-to postgain measures that we have been using because of the comparative nature of our research) may significantly improve the precision with which we can assess relative learning ability. Or, as suggested by the continuum model, it may be more feasible to use as a second stage in the process of assessing individuals who have already been identified as having possible learning difficulties. The mediation procedure on the other hand may prove to be better suited as a diagnostic/prescriptive device in view of what appears to be a relatively greater emphasis on contingency in instruction and metacognitive skills.

With this as background, we can consider suggestions for directions of future research. Generally-speaking, further research is warranted on the concurrent and predictive validity of the measures, and in view of our previous comments about prediction, it would seem to make most sense to focus efforts on the graduated prompting procedure.

One avenue that seems important to pursue is the relation between dynamic measures in the same task domain (given the evidence amassed thus far, the perceptual performance domain is the most logical initial choice). We have investigated the relation between dynamic assessment and static transfer and between dynamic assessment and spontaneous learning but not between prompting on a predictor (i.e., SDT) and prompting on a transfer (i.e., Animal Stencils) task. We would hope of course that a rank ordering of children on the basis of their responsiveness to instruction would be similar across tasks.

As suggested earlier, an aspect that should be given consideration is that of using a measure of number of prompts to index learning rather than, or in addition to, a measure of gain. We are in the process of deriving this measure from the research we had already completed, and may well find that it represents a more sensitive estimate. Brown and colleague who routinely use this measure, have discussed the merits of the approach (see Brown & Ferrara, 1980), not the least of which is that it enables one to discriminate between children who show similar gains but who differ in the amount of instruction needed to make these gains.

The above mentioned recommendations follow from studies already completed. Let us now consider the research agenda from a broader perspective: It seems clear we need to investigate other types of criterion measures in our validation studies. For example, we need to look at the extent to which dynamic measures can predict to classroom tasks, to learning in classroom environments and to learning in response to interventions generated by the assessment. Ultimately, we would want to do this longitudinally as well as concurrently. While field research seems somewhat premature at this point, one way to begin to address these issues might be to undertake instruction of a curriculum unit under laboratory conditions and look at the relation between dynamic measures and various criterion referenced measures.

Another issue which has no doubt occurred to the reader is that the research conducted thus far needs to be extended to other task domains; our research has focused on perceptual performance and to a lesser extent on quantification, although work on developing dynamic procedures and criterion and predictor tasks in the verbal domain is ongoing. In addition, the research needs to be extended to older children, though again, we have begun to develop procedures and tasks for elementary age children.

5.0 Preliminary Findings on Dynamic Assessment and Educational Prescriptions

An issue of major concern to us has been the utility of dynamic procedures, in particular mediation, for deriving prescriptive information. The need for such measures goes undisputed. The available diagnostic tests are not very well standardized (Arter & Jenkins, 1979), nor do they provide information about a child's learning processes, that is, information about task-specific and general learning strategies or about remedial strategies. Dynamic procedures on the other hand would seem to be ideally suited for such purposes. Indeed, Feuerstein's LPAD is in large part used to diagnose cognitive functioning.

5.1 Observations of Young Children's Learning

Our work on prescription has focused on observational measures of young children's general cognitive strategies. The long-range goals of this work are to identify a set of behaviors that can be coded reliably and which reflect cognitive strategies determined to underly successful performance, and to identify instructional methods that will foster the development of such strategies (and eliminate the use of ineffective ones).

We should make it clear that our emphasis on general cognitive strategies does not imply that we believe that knowledge base and task-specific strategies are unimportant. To the contrary, we view assessment of these skills as equally critical. Out of necessity, however, we had to initially narrow our research focus, and both previous research and the age of our target population led us to choose general strategies as a starting point. Age entered into our decision in that we reasoned that knowledge base would be less of a factor for young children on the tasks we had selected than it would be for older children on school-like tasks. The use of an observational methodology does not imply an exclusive commitment. In fact, we are in the process of developing a scheme for analyzing the errors young children make during mediation. Hopefully, this will provide further information about "bugs" in a child's thinking.

The starting point for our studies on prescription was an investigation by Burns and colleagues (Burns et al., 1985). Her study involved observing four- and five-year old children as they performed a series of tasks (including the SDT). Observations were coded using the behavior categories contained in Table 3. A subset of these (the six categories marked with

---

Insert Table 3 about here

---

asterisks) were found to discriminate high from low test scores and/or the performance of low SES from high SES children. For four of the categories (the exceptions being Information Giving and Visual Scan) high frequencies were associated with poor and/or low SES performance. For the two exceptions, low frequencies were associated with poor and/or low SES performance.

The nature of the strategies presumed to underlie these categories deserves comment. In developing the coding scheme, Burns was influenced by Feuerstein's propositions concerning deficient cognitive functioning, as well as by other cognitive models (Brown, Bransford, Ferrara & Campione, 1984; Brown & DeLoache, 1978; Bransford & Stein, 1984; Flavell, 1979) emphasizing the importance of metacognitive skills such as planning and monitoring . For example, the Visual Scan, Looking at the Model, and Inappropriate Manipulation of Materials categories could be considered to be behavioral representations of Feuerstein's deficient cognitive functions of unsystematic search, lack of comparative behavior and trial and error behavior, respectively. Other categories, such as Used All Blocks, Information Giving and Corrects Self appear to capture problem definition, monitoring and planning strategies.

In all our research on dynamic assessment, we have videotaped the experimental sessions. These are then coded to derive measures on the behavioral categories described above. (It also provides a record of tester behavior which is coded to insure that the tester has adhered to the prescribed dynamic procedure). We will discuss the results of two of our studies, Study B and a single-subject project, as they relate to prescription.

In Study B, handicapped and non-handicapped children were assigned to either a mediation, graduated prompting or static (i.e., task demonstration only) assessment condition. For part of the study, children received the following sequence of treatments: 1. pretesting on the SDT, 2. mediation, graduated prompting or static assessment on the SDT, 3. posttesting on the SDT. Observational data are available for all sessions, and thus far have been coded for the Corrects Self categories.

Our purpose in collecting these data was to investigate the degree to which dynamic assessment was helpful in teaching general strategies that facilitate performance. We wished to know whether our interventions could positively influence strategy use, in which case this might be usefully communicated to teachers. Figure 5 depicts the changes observed in

Insert Figure 5 about here

self-correction. As can be seen, dynamic assessment is generally effective for creating change in strategy use for the SDT.

While the results of Study B are encouraging, it is nevertheless clear that some handicapped children do not benefit from our interventions. This of course makes it difficult to specify prescriptive information that might help teachers help these students learn more effectively. To address this issue, we undertook a series of longer-term assessments. One of these was a single-subject project whose performance data was mentioned earlier. What was not discussed at that point was the nature of the mediation that we provided. After the first sessions of mediation, we began to tailor the intervention to address particular behavior categories. Using both behavior category frequencies from the first several sessions and examiner judgements, we identified a strategy that appeared relatively problematic for the child. Following this, all subsequent sessions emphasized remediation of the targeted behavior in addition to the standard mediation interventions. Low frequencies of visual scanning and self-correction were identified in Bobby and Adam, respectively (see Figure 6).

Insert Figure 6 about here

Consider the case of Bobby. Th_ child was observed to search only a portion of the stencils before selecting one. To overcome the problem, the examiner first taught the child to use his "eyes" to look at all of the stencils. The examiner would model the behavior of using her eyes to look at each stencil. The looking was done in a much exaggerated way, and the child was encouraged to imitate. The examiner next established the rule that the child was required to place his hands at the table's edge and to find the correct stencil using his "eyes" before selecting it. As the rule became more automatic and the frequency of scanning increased, the examiner occasionally used the verbal reminder to "first find it with your eyes". Of course, throughout the sessions the child received praise whenever he scanned appropriately.

Note what happened in the case of Bobby whose tailored mediation commenced on trial 5. Referring to Figure 6, we can see that the examiner's interventions resulted in increased visual scanning on the SDT and on the transfer tasks as well. The improvement on transfer, though not as dramatic as that observed on the SDT, is nonetheless remarkable since the examiner did not intervene during performance on these tasks. Notice also that at the point where the examiner begins to focus on visual scanning behavior is the point at which he begins to evidence some consistency in improved performance on the SDT. In view of this, it seems worthwhile to consider modifying future designs in such a way as to allow determination of whether the observed improvements in performance are due to the tailoring, the standard mediation or both. For example, for a child such as Bobby, it would mean using a control child who exhibited relatively little visual scanning behavior and who received only standard mediation.

To summarize, the results suggest that mediation is generally effective for remediating strategies that help children perform effectively. In addition, we have described the strategies we assess and our procedures for assessing them. Finally, we described a model for conducting longer-term assessments that can be used when standard mediation is not sufficient and from which one can derive and test prescriptions that are tailored to the individual child.

4.3 The Utility of Dynamic Assessment of Generating Prescriptions: Directions for Future Research

While our research is at a preliminary stage, it seems clear that even a relatively brief session of dynamic assessment can provide valuable insights into a child's cognitive processing. In our work with young children we have focused on the assessment of general cognitive strategies that reflect planning, monitoring, and so forth, while with older children -- we regret that space does not permit detailed discussion of our work with these children -- we have been more concerned with task-specific knowledge and strategies. Any prescriptive assessment will ultimately need to address all three aspects.

Another important aspect of research on prescription is the method by which the prescriptions are derived, and our work illustrates a number of the available options: observation, interaction, and error analysis. In each case, the requirement is that the method can reliably produce valid prescriptions. While our work is just beginning, we feel we have made progress in this direction. For example, the observational measures that we have used with young children are easily derived, and the results of Burns et al.'s (1985) study, and our single-subject project indicate that the behaviors are closely associated with task performance. We also suggested a methodology for establishing a more direct causal link between behavior and performance, and which could be easily applied to each of the behavior categories.

There are several other avenues of research that could be pursued. One that seems important is to establish the validity of the behavior categories for other tasks, in particular tasks that a child might receive in school. A second, perhaps more basic, avenue would be to gather convergent and discriminant validity data on the strategies assumed to underly the behavior categories.

Another important aspect of a prescriptive assessment is the nature of the instruction given to children. Our initial attempts at developing mediation for older children clearly demonstrated that the method of instruction influences the richness of the information that is generated about children's learning. Similarly, it appears from Study B that standard mediation, more so than graduated prompting, is an effective technique for changing strategy use.

It is a delicate balance indeed. It is necessary to find an instructional approach that makes the child's cognitive processes as explicit as possible and at the same time proves effective for remediation so that this information can be communicated to teachers.

5.0 Factors Related to the Implementation of Dynamic Assessment

Teachers are the ultimate recipients of information derived from dynamic assessment. For this reason we have been concerned with investigating a number of issues related to implementation: the association between observation of dynamic assessment and a change in a teachers' perceptions about a children's abilities and potentials, teachers' evaluations of the utility of information provided by dynamic assessment, and the extent of interactions between judgments about utility and instructional orientation.

Overall, a review of the literature suggests two major conclusions. First, it appears that expectations can play an important role in a teacher's prescriptive planning for a child, and second, that seeing positive change in a child's performance can help alter initially low expectations. These data support the idea that one important role of dynamic assessment may be to change the pessimistic attitudes of teachers toward handicapped children and to convince them that the children have more potential to learn than is traditionally thought.

In our first study, we compared the responses of teachers who viewed an interpreted segment of a standard, static assessment session with their responses following viewing of an interpreted segment of a dynamic assessment session on the same child. Assessment sessions were presented to the teachers via an interactive videotape system. Introductory material was the same across both conditions. Each taped segment was divided into several parts. Clarifying information (e.g., information that the child chose the correct solid card but the wrong stencil) was presented on the computer screen after each trial.

There were two treatment groups. Both groups saw one segment of a child participating in a static assessment session. Group 1 then saw a second segment of the same child participating in another static assessment session. Group 2 saw a second segment of the same child participating in a dynamic assessment session. After viewing each assessment segment (either static or dynamic), each participant completed an 18-item questionnaire. The items on the questionnaire were grouped into three subscales: (1) Task Involvement (including items such as "Was the child attentive, persistent, interested in doing well?"); (2) Task Specific Knowledge and Strategies (with questions like "Did the child know the names of relevant shapes, look at all the materials, compare his work to the model?"); and (3) General Competence ("Was the child competent, successful, aware of his success/failure?").

In Figure 7, we see that all teachers viewed the children as moderately

Insert Figure 7 about here

involved in the task in both the static and dynamic conditions (the ratings hover around 3, the midpoint of the scale), with a small decrease in the ratings from Trial 1 to Trial 2 in the group who saw two static sessions.

With regard to Task Specific Knowledge and Strategies (Figure 8), all

Insert Figure 8 about here

teachers considered the children to be somewhat low (about 2.5) after viewing the first segment of static assessment, but those teachers who then viewed the same children during dynamic administration of the same task rated them much higher (about 3.5) than did those who viewed an additional segment of static assessment.

In Figure 9, we see an even more dramatic shift in judgements of the

Insert Figure 9 about here

children's General Competence (from 2.1 to 3.8) by the teachers who viewed dynamic assessment segments during Trial 2, as contrasted with those who remained firm in their low estimations of the children's competence after viewing two brief static assessment sessions.

Overall, then, teachers considered the child they observed to be using more task appropriate strategies and to be generally more competent when viewed in a dynamic testing situation than when viewed in a static assessment session working on the same task. These effects held regardless of the level of training and experience of the teacher, though less experienced teachers did tend to rate the children more positively across trials and assessment types.

The results we have just presented were based on the combined responses of 60 teachers, each of whom saw 1 of 2 children. We will now consider the responses to each of the children separately.

The videotapes used as the stimuli in this study depicted two very different children. It is instructive to compare the responses of teachers to the performance of each of these children on the 2 scales that showed significant effects of dynamic assessment: (a) Task Specific Knowledge and Strategies; and, (b) General Competence.

Frank is 6 years old. He has been classified as mentally retarded/physically handicapped and attends a special school for multiply handicapped children. He scored in the mentally retarded range on the GCI of the McCarthy and his MA was estimated to be 2 years, 10 months. He has poor muscle tone and just began walking alone this year. His speech has been generally limited to one and two word utterances.

Gary is 5 years, 3 months old. He has been classified as mentally retarded/emotionally disturbed and attends a special education school. He also scored in the mentally retarded range on the GCI of the McCarthy and his MA was estimated at 3 years, 0 months. He is strong and big for his age. His speech is often difficult to understand. To summarize, Frank is older, smaller, and more retarded (based on intelligence test scores) than Gary.

All teachers rated Frank as relatively low (around 3) with regard to task specific knowledge and strategy use after the first Trial. Those who saw him in a second static session rated him somewhat higher, but those who viewed the dynamic session ranked him two full, scale points higher than they had after Trial 1 (see Figure 10). Teachers rated Gary higher than they did Frank after

Insert Figure 10 about here

Trial 1, accurately reflecting his higher level of functicning as measured by the McCarthy. After Trial 2, teachers who saw more static assessment rated him significantly lower than they had earlier, while teachers who viewed him during dynamic assessment made no change in their rating (see Figure 11). It

Insert Figure 11 about here

appears that the nature of this child's emotional disturbance may be a factor in the lower expectations of his performance on cognitive tasks as a function of getting to know him better -- his poor prognosis as a learner may be perceived more as a function of his emotional disturbance than his cognitive ability.

The same pattern is repeated on the General Competence Scale, with Frank being rated very low after Trial 1, but very high after the dynamic assessment session of Trial 2, an actual increase of 2.4 scale points. On the same scale, Gary was again rated higher than Frank on first impression. On Trial 2 the static group rated his performance significantly lower, while the dynamic group judged him to be significantly more competent.

In summary, teachers initial ratings of both children coincided with their relative standing as measured by a normative measure; Frank was viewed as less strategic and less competent than Gary. In addition, there were differential effects of static versus dynamic assessments for both children on both scales, all group differences on Trial 2 were statistically significant. The results also indicated that Frank's ratings improved in both testing conditions on Trial 2, but the dynamic assessment always produced far larger rating changes. Finally, Gary's ratings decreased over trials on both the Task Specific and General Competence Scales after repeated static assessment, while they held or increased after dynamic assessment.

What does all of this say about dynamic assessment for young handicapped children? Taken in the context of the literature on the effects of expectations on teachers' behavior towards handicapped children, we believe our data demonstrate the potential benefits of including teachers in the dynamic assessment process. If, as the literature suggests, a teacher's expectation of a student's potential for learning has direct impact on the level and type of effort that teacher devotes to the child, then the effects of dynamic assessment on teachers' judgments of strategic ability and general competence that we have demonstrated here have important implications for the way we should approach assessments of handicapped children.

6.0 Summary and Conclusions

Our research has focused on three major topics: the utility of dynamic assessment for predicting learning, the utility of dynamic assessment for generating educationally relevant prescriptions, and factors associated with the implementation of dynamic procedures. While many interesting and important questions remain to be answered about these topics, our research offers several conclusions and suggestions for future research.

Consistent with earlier reports, our work indicates that static and dynamic measures do indeed produce different estimates of learning. A substantial number of our children learned to perform tasks when static measures suggested that they would do otherwise. Furthermore, we have demonstrated that this discrepancy is not an artifact of using fullscale rather than subscale static measures. Nor is the discrepancy likely due to any unreliability associated with the measures themselves.

The validity of dynamic measures is indicated by data from several experiments. It appears that for handicapped and non-handicapped children alike, responsiveness to instruction is predictive of within-domain transfer performance. On the other hand there does not seem to be a strong relation between responsiveness to instruction across domains, suggesting that dynamic measures, if and when implemented, would need to be taken in each domain for which one wanted to predict learning.

In sum it appears from research conducted thus far that dynamic assessment is useful for predicting learning. It seems clear however that we need to consider other criteria in order to broaden the base of our validation studies. For example, we need to look at the extent to which dynamic measures can predict to classroom tasks and to learning in classroom environments. And, of course, we want ultimately to do this longitudinally as well as concurrently.

Our research on the topic of prescription has indicated that mediation assessment is generally effective for remediating strategies that help children improve their performance. Moreover, we are particularly encouraged by our experiences with extended mediation. Using observational information, we have tailored our mediation to the needs of individual children which in turn has produced changes in the targeted strategies and improvements in task performance. The model of extended mediation is promising because it allows us to derive information about ineffective learning processes and to observe the results of our attempts to remediate these processes. Such information may ultimately be of value to teachers. However, research must begin to address the issue of the generality of the behavior categories, in particular for tasks a child might receive in school. As discussed earlier, future research also needs to consider the assessment of task-specific knowledge and strategies in addition to general cognitive strategies.

Finally, our research investigating the impact of implementing dynamic procedures illustrate that dynamic assessments can have important effects on teacher expectations. While encouraged by this, we are nevertheless aware that changes in expectations will not necessarily produce changes in teaching behaviors, especially if teachers do not have a better idea of how to teach in ways that help various children learn. This is one of the reasons why we are placing more and more emphasis on using assessments to provide prescriptions for teaching.

We are greatly encouraged by the potential of dynamic procedures and their future role in psychoeducational assessment.

References

Arter, J. A., & Jenkins, J. R. (1979). Differential diagnosis - prescriptive teaching: A critical appraisal. Review of Educational Research, 49(4), 517-555.

Arthur, G. (1947). A point scale of performance tests. New York: Psychological Corporation.

Belmont, J. M., & Butterfield, E. C. (1979). The instructional approach to developmental cognitive research. In R. V. Kail, Jr. & J. W. Hagen (Eds.), Perspectives on the development of memory and cognition. Hillsdale, NJ: Lawrence Erlbaum Associates.

Bransford, J. D., & Stein, B. S. (1984). The IDEAL problem solver. New York: W. H. Freeman and Company.

Brown, A. L., Bransford, J. D., Ferrara, R. A., & Campione, J. C. (1984). Learning, remembering, and understanding. In P. H. Mussen (Ed.), Handbook of Child Psychology. New York: John Wiley & Sons.

Brown, A. L., & DeLoache, J. S. (1978). Skills, plans, and self-regulation. In R. Siegler (Ed.), Children's thinking: What develops? Hillsdale, NJ: Lawrence Erlbaum Associates.

Brown, A. L., & Ferrara, R. A. (1980, October). Diagnosing zones of proximal development: An alternative to standardized testing. Paper presented at Conference on Culture, Communication, and Cognition: Vygotskian Perspectives, Center for Psychosocial Studies, Chicago, Illinois.

Brown, A. L., & French, L. (1979). The zone of potential development: Implications for intelligence testing in the year 2000. Intelligence. 3, 255-273.

Bryant, N. R. (1982). Preschool children's learning and transfer of matrices problems: A study of proximal development. Unpublished master's thesis, University of Illinois.

Budoff, M. (1967). Learning potential am institutionalized young adult retardates. American Journal of Mental Deficiency, 72, 404-411.

Burns, M. S. (1985). Comparison of "Graduated Prompt" and "Mediational" dynamic assessment and static assessment with young children. Unpublished manuscript, Vanderbilt University, Nashville, Tennesee.

Burns, M. S. Haywood, H. C., Delclos, V. R., & Siewert, L. (1985). Young children's problem solving strategies: An observational study. Unpublished manuscript, Vanderbilt University, Nashville, Tennesee.

Campione, J. C., Brown, A. L., & Ferrara, R. A. (1982). Mental retardation and intelligence. In R. J. Sternberg (Ed.), Handbook of human intelligence. New York: Cambridge University Press.

Campione, J. D., Brown, A. L., Ferrara, R. A., Jones, R. S., & Steinberg, E. (1983). Differences between retarded and nonretarded children in transfer following equivalent learning performance: Breakdowns in flexible use of information. Unpublished manuscript, University of Illinois, Champaign-Urbana, Illinois.

Feuerstein, R., Rand, Y., & Hoffman, M. (1979). The dynamic assessment of retarded performers: The learning potential assessment device, theory, instruments, and techniques. Baltimore: University Park Press.

Flavell, J. H. (1979). Metacognition and cognitive monitoring. American Psychologist, 34, 906-911.

Lidz, C. S. (1983). Dynamic assessment and the preschool child. Journal of Psychoeducational Assessment, 1, 59-72.

McCarthy, D. (1972). Manual for the McCarthy Scales of Children's Abilities. New York: The Psychological Corporation.

Reynolds, C. R., & Clark, J. H. (1983). Assessment of cognitive abilities. In K. Paget & B. Bracken (Eds.), The psychoeducational assessment of preschool children. New York: Grune and Stratton.

Simner, M. L. (1983). The warning signs of school failure: An update profile of the at-risk kindergarten child. Topics on Early Childhood Special Education, 3, 17-27.

Vygotsky, L. (1962). Thought and language. Cambridge, MA: The MIT Press.

Vygotsky, L. (1978). Mind in society: The development of higher psychological processes. Cambridge, MA: Harvard University Press.

Footnote

[1]Preparation of the chapter was supported by Grant G0083C0052 awarded to Vanderbilt University by the U.S. Department of Education, Office of Special Education, and by Grant T238A1 awarded to the first author by the Natural Sciences and Engineering Research Council of Canada. The authors would like to acknowledge the other project investigators, Carl Haywood and Ted Hasselbring, and our research assistants, Keith Allred, Ron Buen, Randi Glorski, Judy Johnson, Stan Kulewicz, Tamra Ogan, Kim Sloan, Deborah Stephens, and Julie Tapp. Finally, we wish to thank all of the children who participated in this research.

Appendix A

Mediation Procedure for the Stencil Design Test

Familiarizing the Child with Materials and Relevant Dimensions

1. Point out cut-outs (I CUT THEM OUT).

2. Label shapes, If there is resistance or difficulty learning labels, then tell the child the label, but go quickly to finding shapes that match and say FIND ALL THE CARDS LIKE THIS. Comment on the lack of labels in a report, but do not get bogged down--the matching encourages comparative behavior while establishing shape as a relevant feature.

3. Point out solids (NOT CUT-OUTS--NO HOLES). Note all are in bottom row, right near child.

4. Label colors (see notes for label shapes).

5. Have child COUNT THE SOLIDS. Focus here is not on the ability to count, but on the child's conceptualization of "solid" or "not cut-out". If child counts correctly to 6, then the distinction is being made.

   * * * WARNING * * *

   If child cannot count all the solids, you need more work on the preceding concepts.

6. Compare 2 circles (big, small)/ 2 white squares (straight, crooked)/ 2 blue cards (solid, cut-out)/ 2 yellow cards (solid, cut-out)/ yellow and blue crosses (yellow, blue).

7. At some point, put solid and cut-out back in wrong place--again to gauge whether the discrimination is being made.

   * * * WARNING * * *

   If child cannot see that you put solid back in wrong place, you need more work on preceding concepts.

Combination Rules

1. Demonstrate what happens when a green circle is placed on a yellow solid. Point out <u>2</u> colors, made from 1 ± 1.

2. Change solids, showing that the <u>inside</u> color changes by changing <u>solids</u>. Allow child to try 1 or 2 color changes. Emphasize that it is <u>solid</u> that is changing.

   * * * WARNING * * *

   If child cannot change the color of the solid, you need more work on the preceding concepts.

Appendix A (cont'd)

3. Use white solid with green circle. Change cut-outs (don't reproduce any of the upcoming designs). Show that outside color changes by changing cut-outs.

4. Put solid on top of cut-out and establish necessary order rule and reason. Have child repeat the rule "I put a cut-out on top of a solid and the color of the solid is in the middle."

5. End with the sample design formed from stencils, then introduce the sample design model.

Helping the Child Reproduce the Model

1. Display model while reproduction is still on the table, discussing how a picture was made of it. Point out that there are 2 colors in the picture and 2 colors in the reproduction, but only 1 color on each separate card.

2. Put stencils back in place and request reproduction. Teach search pattern over cut-outs and over solids. Have c[illegible] say "Is it this one?".

3. When production is made, encourage checking back [illegible] model. Go over what is right and what is wrong about the production.

* * * WARNING * * *

If the child's production is wrong, you need more work on the preceding concepts. Refer to any errors made in route to a correct answer (spontaneous corrections) and discuss why they were wrong. Alternate the correct one and the wrong one. Always end with the correct solution.

4. Repeat Step 3 with each of the remaining training models.

Appendix B

Graduated Prompting Procedure for the Stencil Design Task

1. DO YOU REMEMBER HOW YOU DID IT WITH THE LAST ONE? If so, HOW DID YOU DO IT? If not, point out and label the solid cards and the cut-outs, then explain that a solid and a cut-out are put together to make one that looks just like the model.

2. LOOK AT ALL THESE CARDS (point out each card individually; a pencil is useful for this). EVERYTHING YOU NEED TO MAKE THIS ONE IS HERE. SEE IF YOU CAN MAKE ONE THAT LOOKS JUST LIKE THIS ONE.

3. SEE THIS MODEL (point to the model)? DOES ONE OF THESE (point to stencils) LOOK JUST LIKE THE MODEL? If child responds no, say RIGHT, NONE OF THEM LOOKS JUST LIKE THE MODEL. If the child responds yes, say NO, NONE OF THEM LOOKS JUST LIKE THE MODEL. Then say, YOU SEE IN THE MODEL WE HAVE A (point out and name the color) SOLID AND A (point out and name the color) CUT-OUT. YOU NEED TO PUT SOME OF THESE TOGETHER (point to solids and cut-outs) TO MAKE ONE THAT LOOKS JUST LIKE THE MODEL. SEE IF YOU CAN MAKE ONE THAT LOOKS JUST LIKE THIS MODEL.

4. LET'S LOOK AT THESE AGAIN. THESE ARE THE SOLID COLORS (point). DOES EACH SOLID HAVE ONE COLOR OR TWO COLORS? LOOK AT THIS ONE, FOR EXAMPLE. (Hold up white solid, #5.) If child does not respond correctly, give correct answer.

   THESE ARE THE CUT-OUTS(point). DOES EACH CUT-OUT HAVE ONE COLOR OR TWO COLORS? LOOK AT THIS ONE, FOR EXAMPLE. (Hold up red cut-out, #17.) If child does not respond correctly, give correct answer.

   DOES THE MODEL HAVE ONE COLOR OR TWO COLORS? LOOK AT THIS ONE, FOR EXAMPLE. (Hold up red-over-white sample model). If child does not respond correctly, give correct answer.

   YOU NEED ONE SOLID AND ONE CUT-OUT TO MAKE ONE THAT LOOKS JUST LIKE THE MODEL (point). SEE IF YOU CAN MAKE ONE THAT LOOKS JUST LIKE THIS MODEL (point to item model).

5. LET'S LOOK AT THE MODEL AGAIN. POINT TO (OR NAME THE COLOR OF) THE PART THAT LOOKS LIKE IT COMES FROM A SOLID. Point if the child responds incorrectly. POINT TO (OR NAME THE COLOR OF) THE PART THAT LOOKS LIKE A CUT-OUT. Point if the child responds incorrectly. NOW SEE IF YOU CAN MAKE ME ONE THAT LOOKS JUST LIKE THE MODEL.

6. LOOK AT THIS MODEL. (Show red-over-white sample model.) LET'S SEE WHAT SOLID I NEED TO MAKE THIS ONE. IS IT THIS ONE? Explore the other solids and whether they are correct. LET'S SEE WHAT CUT-OUT I NEED TO MAKE THIS ONE: IS IT THIS ONE? Explore the cut-outs up to the correct one. LOOK AT WHAT HAPPENS WHEN I TAKE A WHITE SOLID AND I PUT A RED CUT-OUT ON TOP OF IT. PART OF THE WHITE SOLID GETS COVERED UP. THAT IS HOW I MAKE ONE JUST LIKE THIS MODEL. (Point to original model.) If the child uses the correct solid, skip prompt 7 and use prompt 8.

Appendix B (cont'd)

7. LOOK AT THIS MODEL. WHICH SOLID COLOR DO YOU NEED TO MAKE THIS MODEL? If the child does not answer say, SHOW ME ON THE MODEL. Demonstrate if child responds incorrectly. THESE ARE THE SOLID COLORS (point). PICK ONE OF THESE. AND SEE IF YOU CAN MAKE ONE THAT LOOKS JUST LIKE THE MODEL.

8. THIS (name the color of the solid) ONE IS PART OF THE MODEL. (Place the correct solid in the center of the board if it is not already there.) LOOK AT THIS PART OF THE MODEL (point to part that looks like a cut-out.) FIND A CUT-OUT FROM HERE (point) THAT LOOKS JUST LIKE THIS PART OF THE MODEL. SEE IF YOU CAN MAKE ME ONE THAT LOOKS JUST LIKE THE MODEL.

9. PUT THIS (name color) CUT-OUT ON YOUR SOLID COLOR. SEE, YOURS LOOKS JUST LIKE MINE.

Table 1

Correlations between McCarthy scores and scores on Stencil Design Task following graduated prompting (N = 44)

	McCarthy Scales			
	General Cognitive Index	Perceptual Performance	Verbal	Quantitative
Stencil Design Task	.18	.48*	-.05	-.15

*p < .01

Table 2

Mean independent and transfer performance (% correct) following dynamic assessment

	Stencil Design	Animal Stencils	Reverse Stencils	*Animal House (X raw score)
GRADUATED PROMPTING (N = 15)	37%	33%	49%	* 18.60
MEDIATION (N = 15)	47%	51%	62%	* 18.33
Children above 75% criterion on Stencil Design Task				
GRADUATED PROMPTING (N = 5)	85%	73%	95%	* 29.80
MEDIATION (N = 4)	75%	75%	78%	* 21.50
Children below 75% criterion on Stencil Design Task				
GRADUATED PROMPTING (N = 10)	13%	14%	26%	* 13.00
MEDIATION (N = 11)	36%	42%	56%	* 17.18

Table 3

Behavioral categories and brief definitions (adapted from Burns et al. (1985))

Attention - looks at experimenter or materials during instructions and/or looks at materials while performing.

Attention & On-Task Manipulation - active contact, using hands, with the materials that the child is working with. This is applicable only when it is time to be manipulating materials.

Off-Task Behavior - active contact, using hands, on the environment or body that is not part of the material in the study. This includes manipulating task materials when the child should be listening to instructions.

*Information Giving - child explains what he/she is going to do before performing the task and/or explains intermediate steps. This information is specific in nature.

*Visual Scan/Looking at Model - looks at model or head moves past the center line (imaginary) dividing the left and right sides of the materials.

*Corrects Self - the child gives an answer and without any intervention from the experimenter, changes the answer.

*Confirmation Seeking, Helpless Gestures & Verbalizations - child looks to the tester while using the task materials or asks for help in a non-specific request.

On-Task Comments - comments made by the child about the task that are not specific to the task completion.

*Inappropriate Manipulation of Stencils - the number of stencils that the child touches that are not a part of the model design that is being made.

Speaking Out Before Instructions Finished - the child speaks, gestures, or starts the task before the instructions are finished.

*Used All Blocks - when making a block design the child uses all nine blocks, even though none of the designs required using all of the blocks.

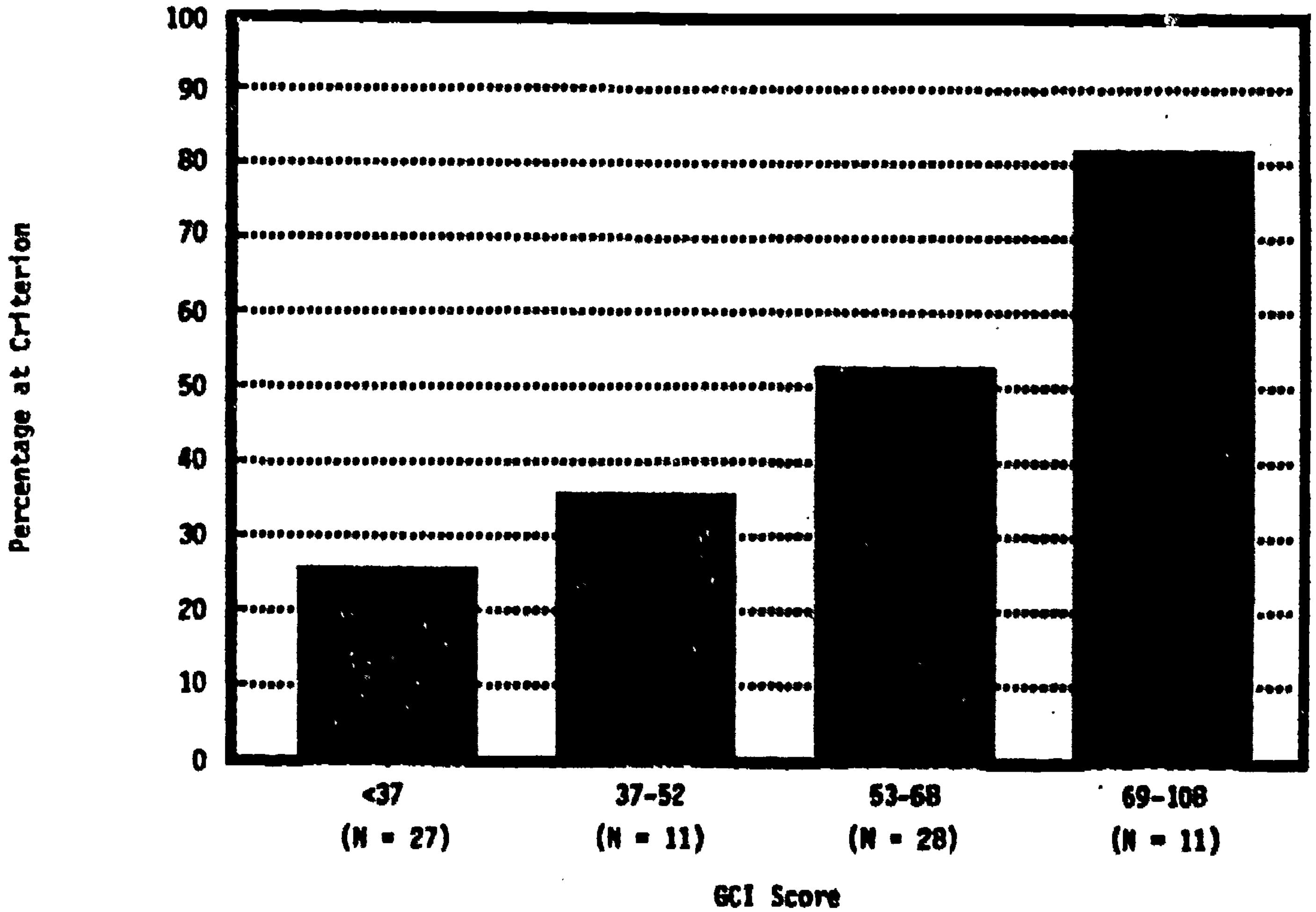

Figure 1. Percentage of children in each GCI group who reach criterion (≥75% correct on SDT).

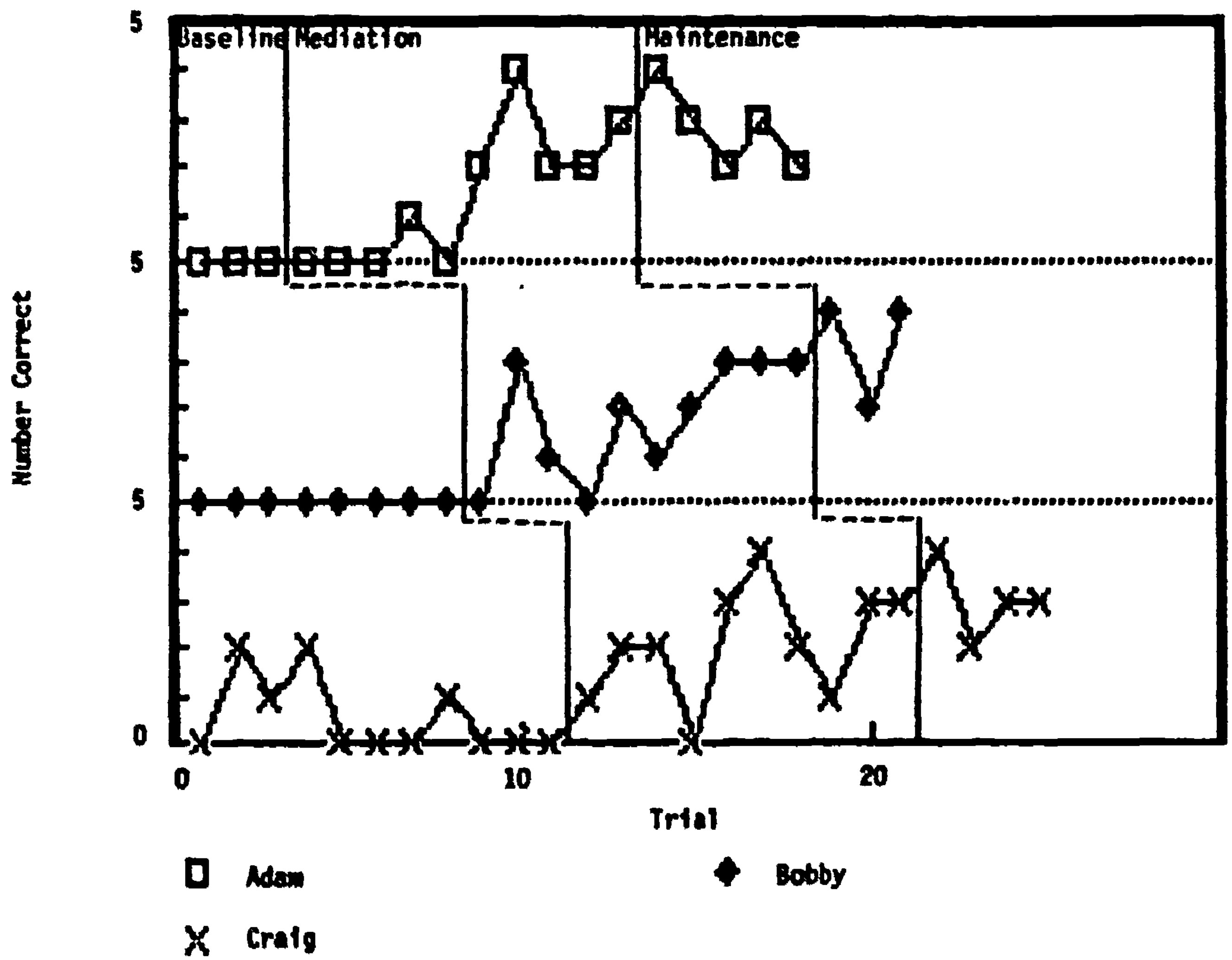

Figure 2. Number of stencils correct on the SDT during baseline, mediation and maintenance trials.

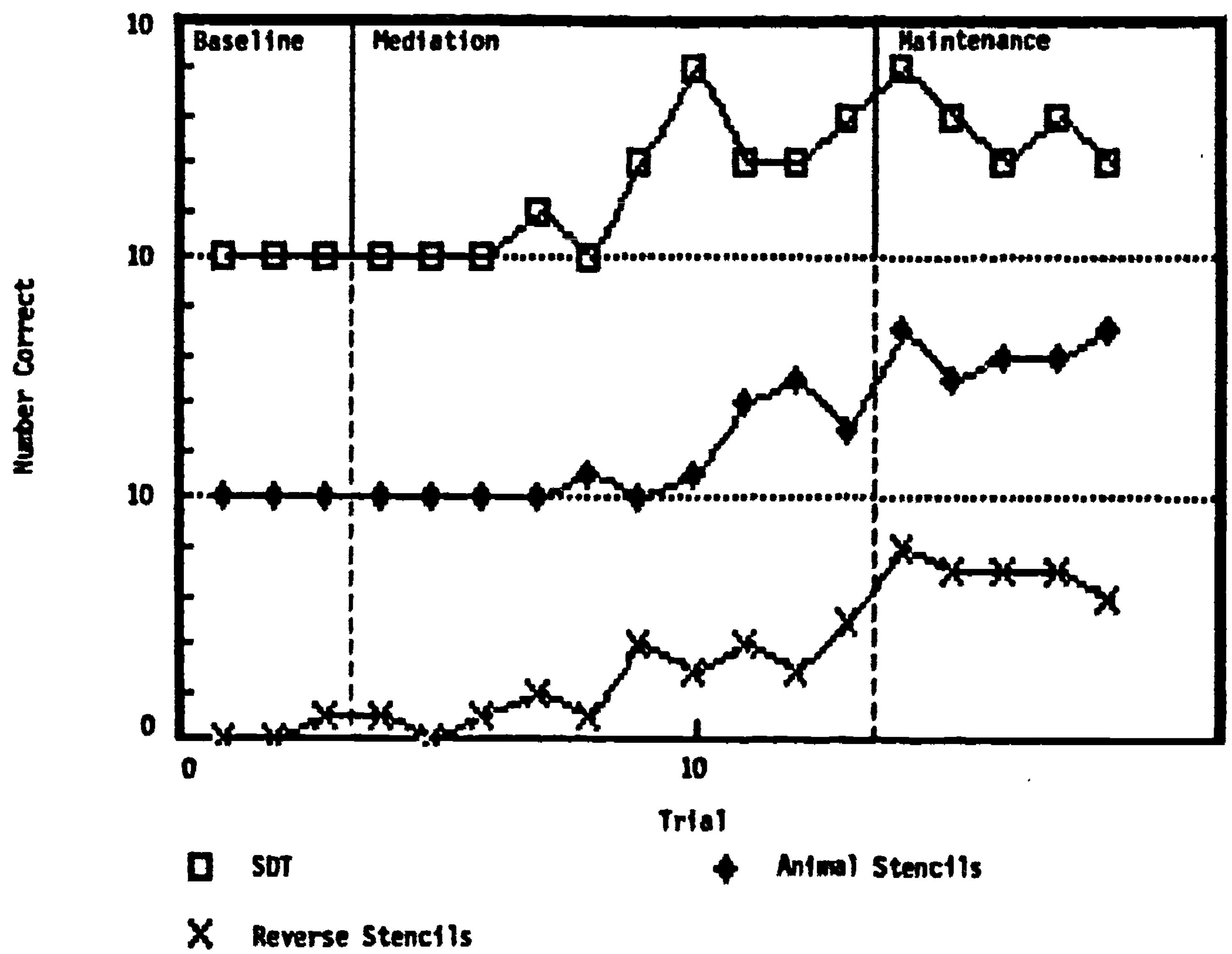

Figure 3. Adam's SDT and within-domain transfer performance during baseline, mediation and maintenance trials.

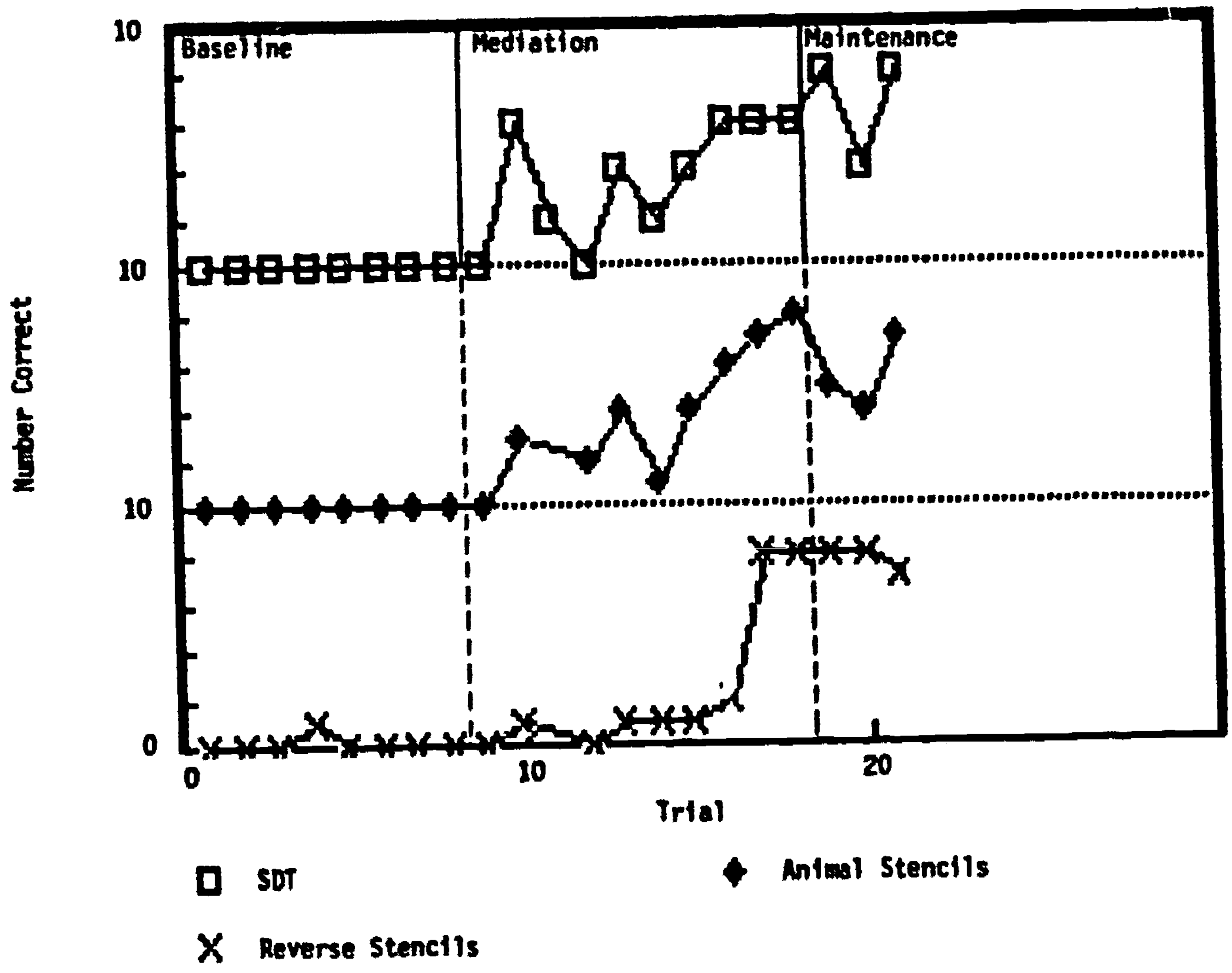

Figure 4. Bobby's SDT and within-domain transfer performance during baseline, mediation and maintenance trials.

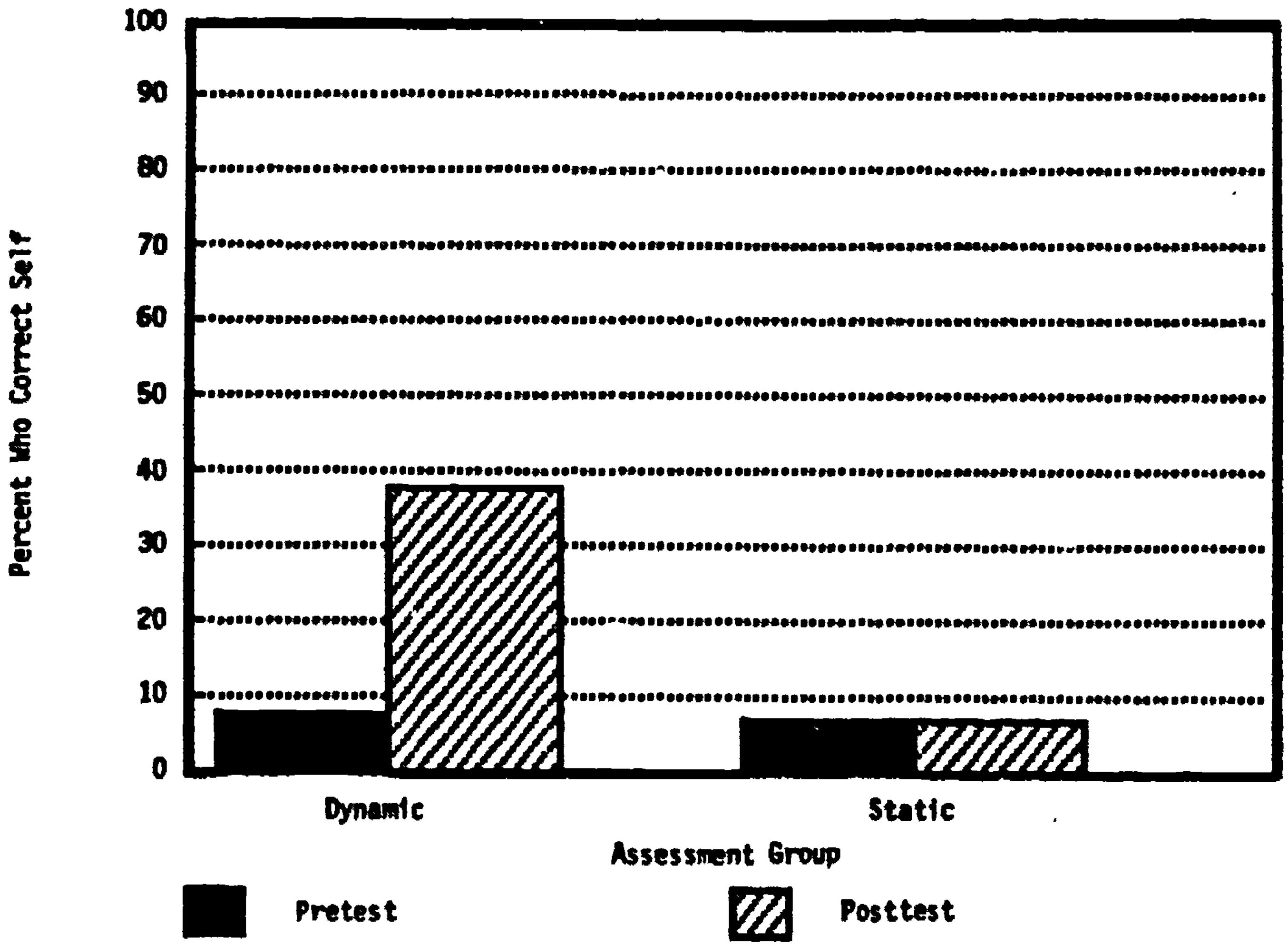

Figure 5. Percentage of Study B children showing self-correction behavior.

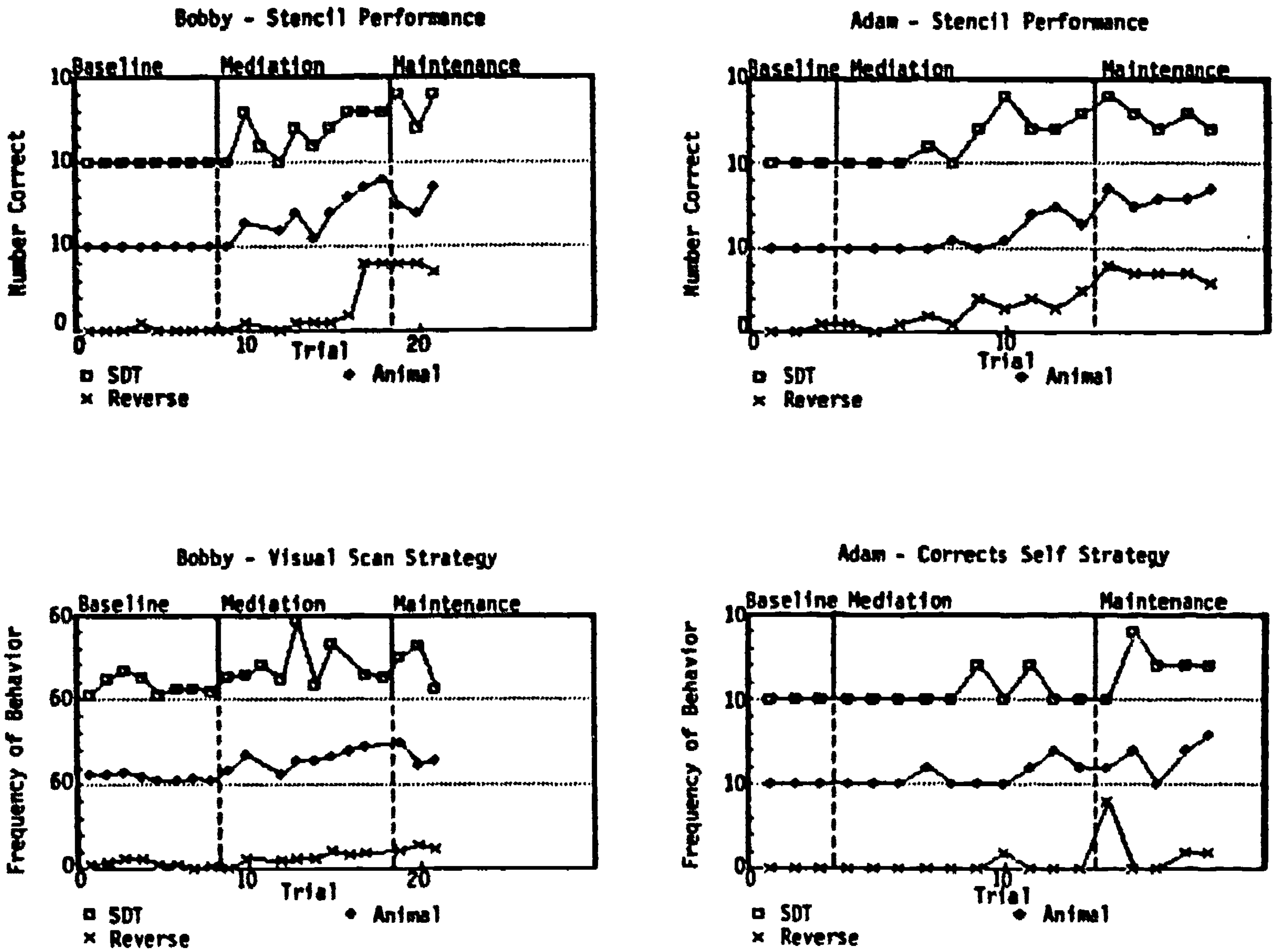

Figure 6. Frequency of strategy use and stencil performance data for Bobby and Adam.

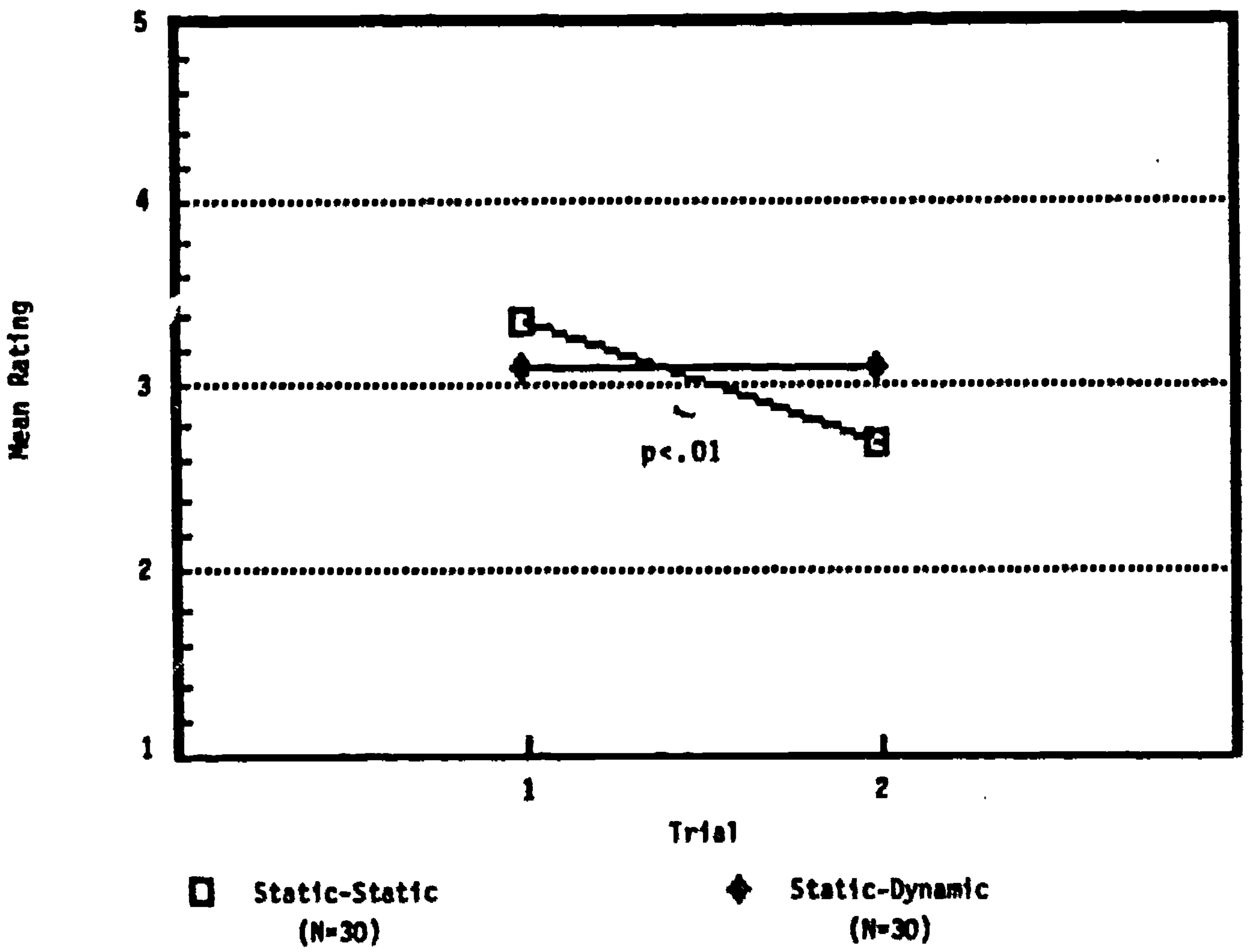

Figure 7. Teachers' ratings of children's task involvement.

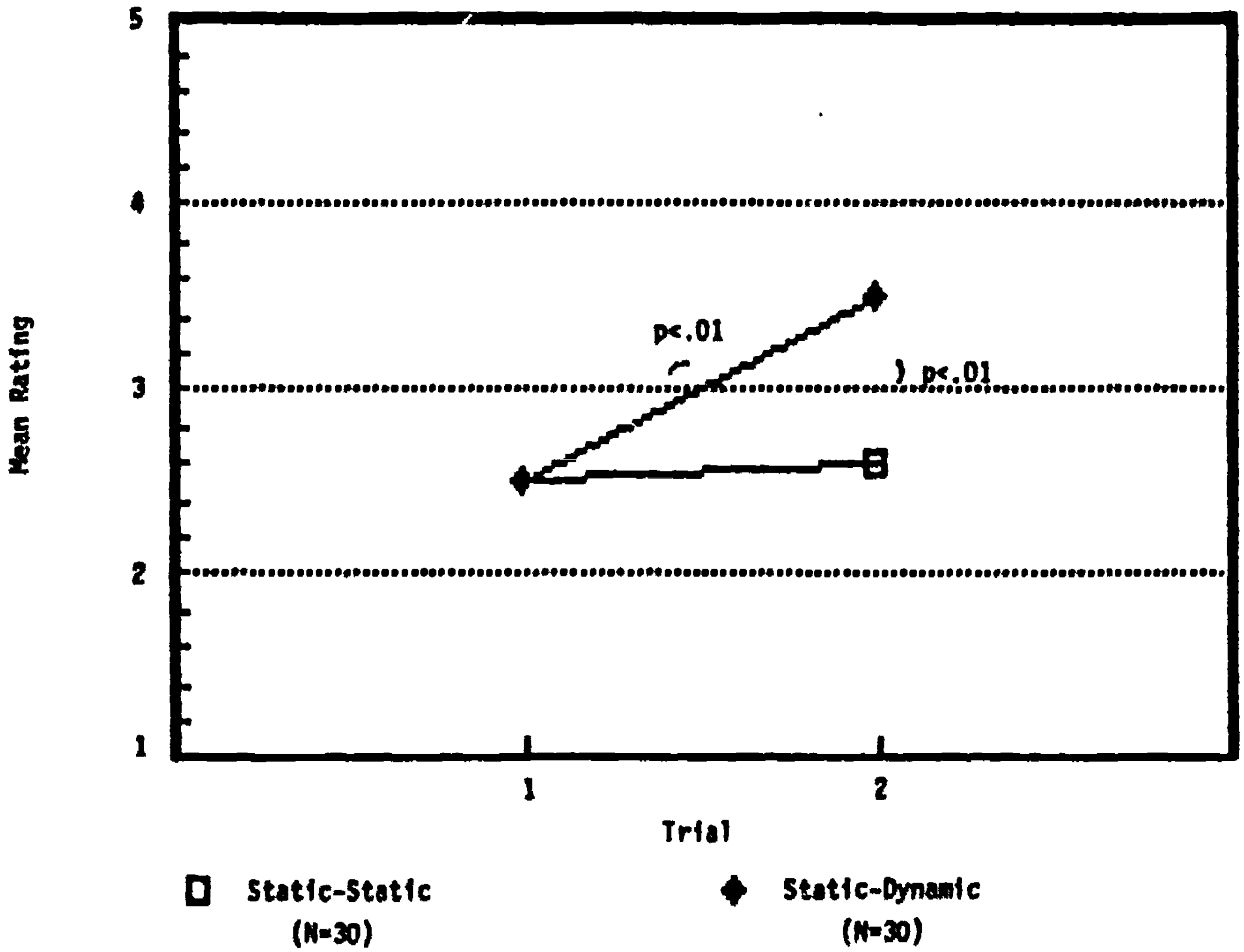

Figure 8. Teachers' ratings of children's task specific knowledge.

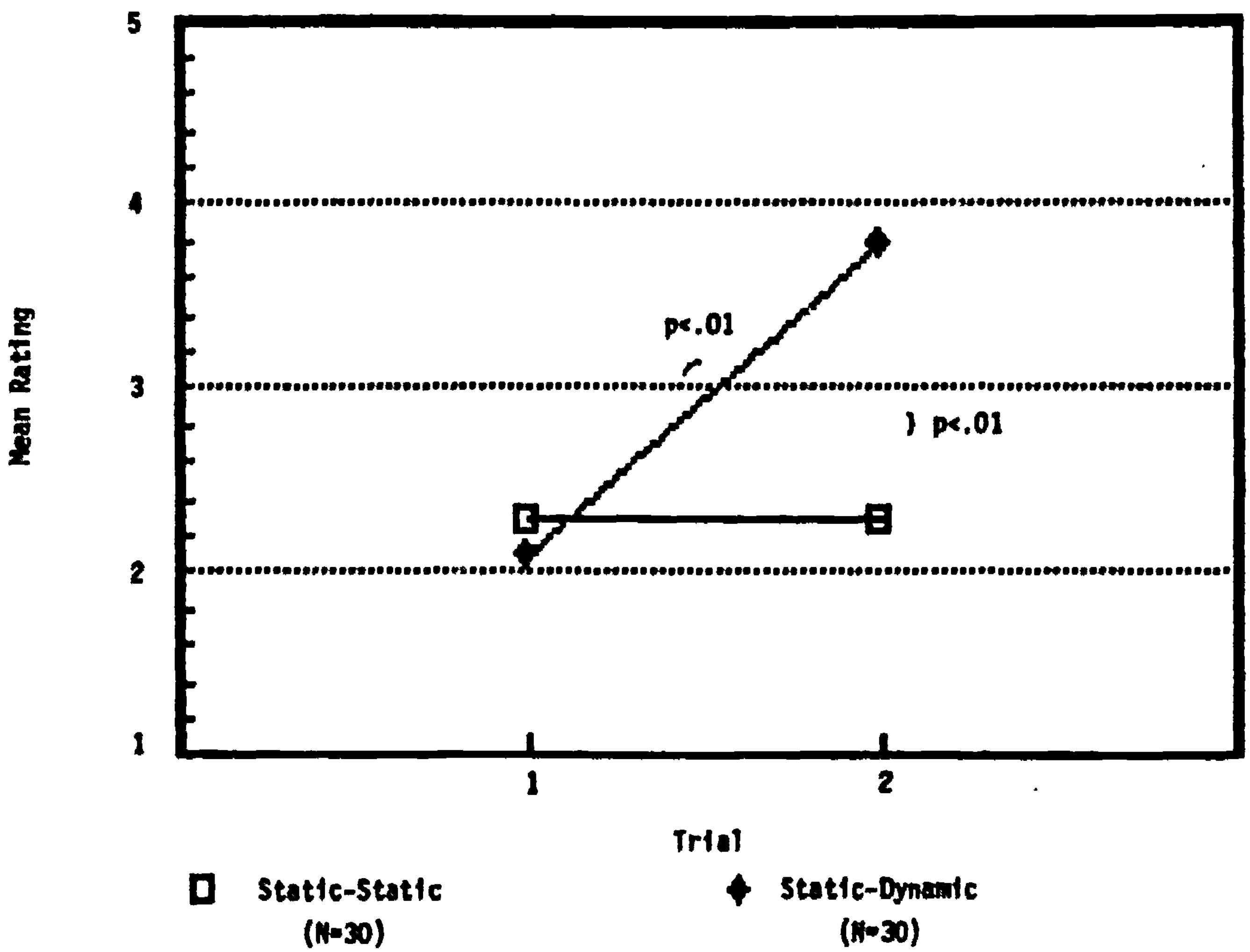

Figure 9. Teachers' ratings of children's general competence.

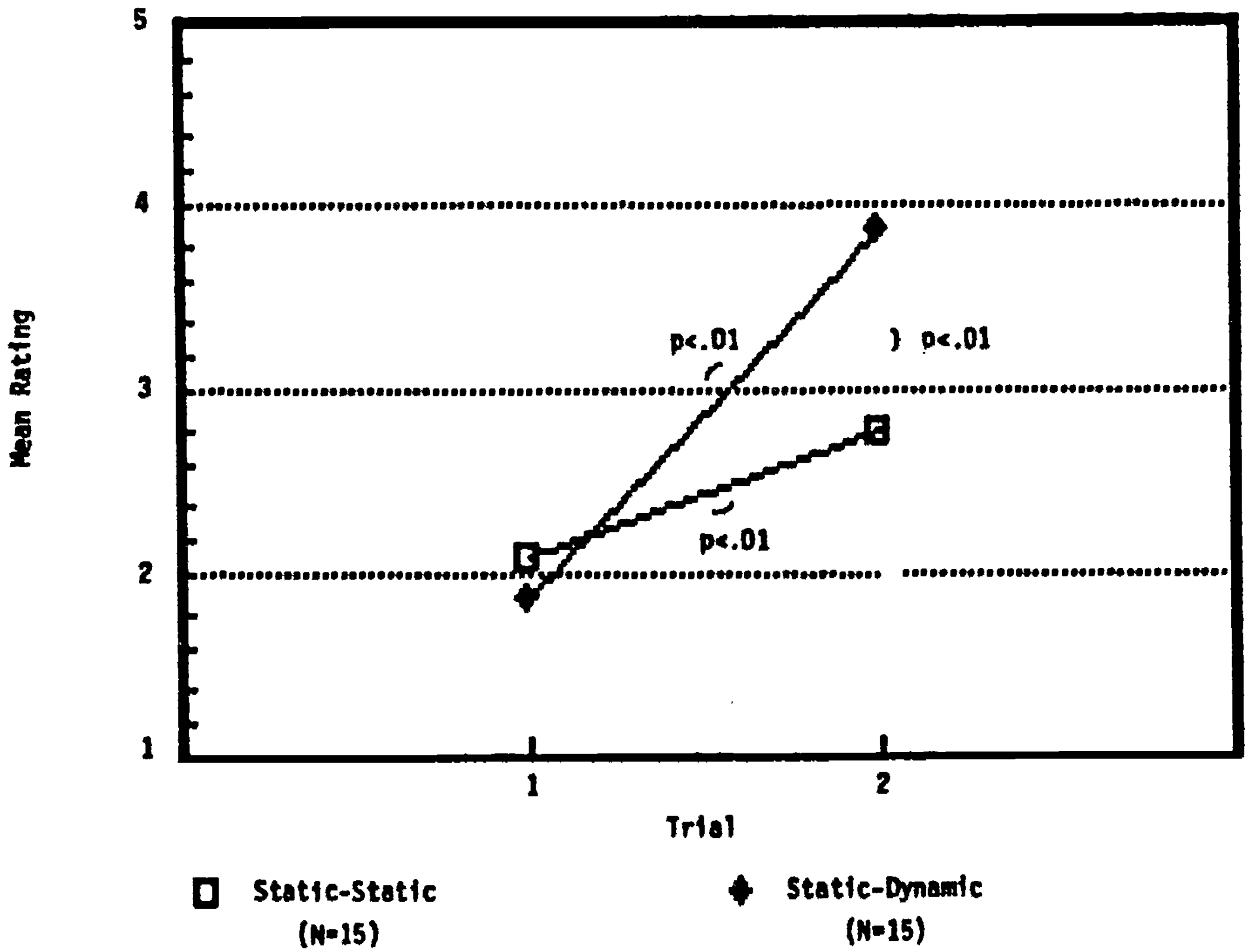

Figure 10. Teachers' ratings of Frank's task specific knowledge.

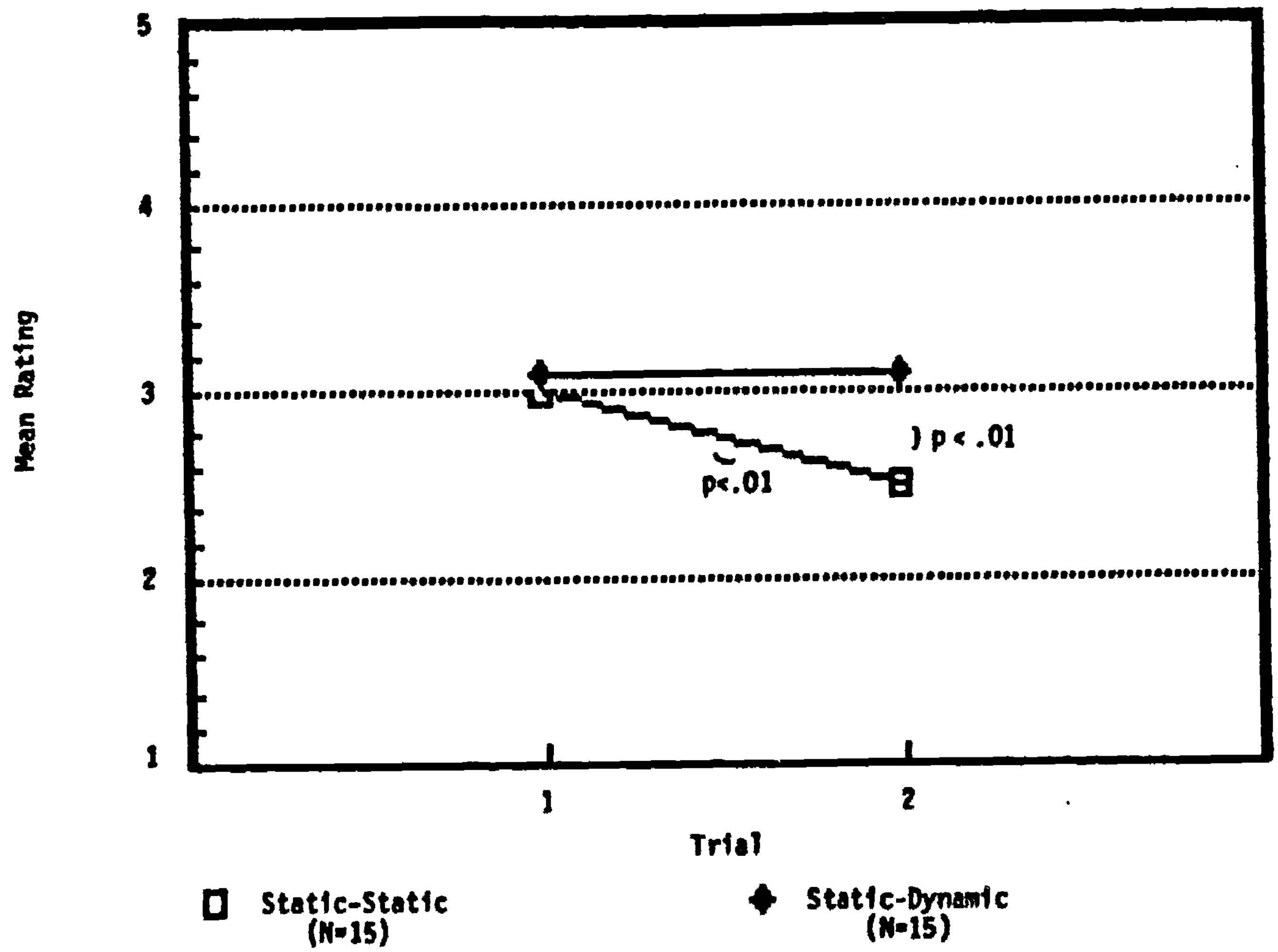

Figure 11. Teachers' ratings of Gary's task specific knowledge.