DOCUMENT RESUME

ED 269 735 CS 008 435

AUTHOR	Campione, Joseph C.; Brown, Ann L.
TITLE	Dynamic Assessment: One Approach and Some Initial Data. Technical Report No. 361.
INSTITUTION	Bolt, Beranek and Newman, Inc., Cambridge, Mass.; Illinois Univ., Urbana. Center for the Study of Reading.
SPONS AGENCY	National Inst. of Child Health and Human Development (NIH), Bethesda, Md.; National Inst. of Education (ED), Washington, DC.
PUB DATE	Dec 85
CONTRACT	400-81-0030
GRANT	HD-05951; HD-06864; HD-15808
NOTE	68p.
PUB TYPE	Reports - Research/Technical (143)
EDRS PRICE	MF01/PC03 Plus Postage.
DESCRIPTORS	*Cognitive Measurement; Cognitive Processes; *Educational Diagnosis; Elementary Education; Identification; Intelligence Quotient; Learning Processes; *Learning Readiness; Learning Strategies; *Predictive Measurement; Predictor Variables; Reading Diagnosis; *Reading Readiness; Reading Research; Standardized Tests; *Transfer of Training
IDENTIFIERS	*Dynamic Assessment; Vygotsky (Lev S)

ABSTRACT

In an effort to validate dynamic assessment methods influenced by Vygotsky's (1978) definition of zones of proximal development (an indicator of readiness), three sets of experiments addressed two goals: the development of diagnostic assessment methods and the use of diagnostic results to guide the design of instructional programs. The first two studies examined two psychological processes presumed to be significant predictors of academic performance: acquiring new information (learning) and using that information in novel situations (transfer). Results indicated the validity of the methods used, showing (1) that groups of children of contrasting abilities do differ in terms of learning and transfer, with differences being greater in transfer; and (2) that learning and transfer are better predictors of performance than either IQ or standardized test scores, with the best predictor being transfer. In the third study, which tested the instructional application of dynamic assessment results (emphasizing enhancing the transfer process), subjects taught to use four reading strategies demonstrated gains in their ability to transfer these skills, as evidenced by their standardized and reading comprehension test scores. Results of these studies suggest that the best predictors of the extent to which individuals are likely to profit from instruction are their initial response to instruction (learning) and, even more, how successfully they can transfer their learning to novel situations. Results also suggest the superiority of dynamic assessment over static testing for planning for instruction. A seven-page list of references is included. (LLZ)

CENTER FOR THE STUDY OF READING

Technical Report No. 361

DYNAMIC ASSESSMENT: ONE APPROACH AND SOME INITIAL DATA

Joseph C. Campione
and Ann L. Brown

University of Illinois at Urbana-Champaign

December 1985

University of Illinois
at Urbana-Champaign
51 Gerty Drive
Champaign, Illinois 61820

Bolt Beranek and Newman Inc.
10 Moulton Street
Cambridge, Massachusetts 02238

The work upon which this publication is based was performed pursuant to Grants HD-05951, HD-06864, and HD-15808 from the National Institute of Child Health and Human Development and Contract No. 400-81-0030 of the National Institute of Education. It does not, however, necessarily reflect the views of these agencies. We would like to acknowledge the considerable input from Nancy Bryant, Roberta Ferrara, and Annemarie Palincsar at various points in the program.

Dynamic Assessment:

One Approach and Some Initial Data

Proponents of dynamic assessment methods are concerned with identifying students who are likely to experience academic problems and with providing descriptions of those students' strengths and weaknesses in such a way that remedial programs can be developed. A major stimulus for the interest in dynamic assessment procedures is a dissatisfaction with certain features of standardized "static" tests. In these static tests, children are asked for specific information or are required to solve certain types of problems. The tester provides no help during the testing session. The score individuals attain represents an estimate of their current, unaided level of competence. All too often, the unwarranted inference is made that these scores are a measure of ability level, i.e., an IQ score of 70 is seen as relatively permanent and resistant to change. In many cases, particularly when children from culturally different backgrounds are involved, this picture may provide a dramatic underestimate of their potential level of performance under more favorable circumstances.

Dynamic assessment methods aim to go beyond this state of affairs by assessing the operation of basic psychological processes presumed responsible for acquisition of the information requested on standard tests. Some children may not have acquired the information or skills being assessed, but nonetheless may be able to do so quite readily if given the opportunity. The future academic performance of children in this category would be expected to be better than one would expect on the basis of their initial, unaided static test performance. To generate this additional diagnostic information, developers of dynamic assessment methods have used a number of different techniques, all of which involve the provision of some form of help to the child. This aid can take the form of modifying the format in which the test is administered (e.g., Carlson & Wiedl, 1978, 1979), providing direct instruction in methods of solving the problems (e.g., Budoff, 1974), or attempting to evaluate directly a set of target processes (e.g., Feuerstein, 1979). The assumption is that performance estimates obtained under these altered conditions will provide more accurate assessments of individual differences than standard test scores, or will at least supplement the picture they paint.

Although sharing common assumptions, the methods that have been advanced differ in a number of ways, including the goal of the program. Some aim to engineer maximal levels of performance; others seek to measure the magnitude of response to instruction; still others focus on the efficiency of operation of specific cognitive processes. Different program goals have resulted in different methods of conducting the assessment. In this chapter, we outline our own variations on the theme of dynamic assessment. The approach we have adopted has been influenced by two lines of research, one specifying the format of the assessment itself and the second identifying the target processes we seek to evaluate.

A General Framework

Our approach to both assessment and instruction has been heavily influenced by Vygotsky (1978) and neo-Vygotskians currently working in the Soviet Union on the development of assessment techniques for recognizing academic delay (Vlasova & Pevzner, 1971; Zabramna, 1971). Both the Soviet investigators and our team have been influenced by Vygotsky's general view of learning and development and his notion of a zone of proximal development. We emphasize, however, that the resultant approach is an amalgam of our views on cognition and instruction and Vygotsky's theory; and it is in no way meant to represent Vygotsky's original views unchanged (Brown & French, 1979; see Mirsch chapter). Vygotsky emphasized that much of learning was mediated through social interactions. Children experience cognitive activities in social situations and come to internalize them gradually over time. At the outset, the child and an adult work on together, with the adult doing most of the work and serving as an expert model. As the child acquires some degree of skill, the adult cedes the child responsibility for part of the job and does less of the work. Gradually, the child takes more of the initiative, and the adult serves primarily to provide support and help when the child experiences problems. Eventually, the child internalizes the initially joint activities and becomes capable of carrying them out independently. At the outset, the adult is the model, critic, and interrogator, leading the child toward expertise; at the end, the child adopts these self-regulation and self-interrogation roles. It is this gradual transfer of control that we seek to capture in ou assessment and instructional sessions.

Within this context, Vygotsky also described the zone of proximal development, which refers to the distance between the level of performance a child can reach unaided and the level of participation that can be accomplished when guided by a more knowledgeable participant. For a certain child, in a particular domain, this zone may be quite small, the interpretation being that the child is not yet ready to participate at a more mature level than his unaided performance would indicate. For another child in that domain, or that child in another domain, the zone of proximal development can be quite large, indicating that with aid, sometimes minimal aid at that, the child can participate much more fully and maturely in the activity than one might suppose on the basis of only unaided performance.

The assessment process suggested by Vygotsky has been quite influential in the diagnostic testing of problem learners in the Soviet Union (Brown & Ferrara, 1985; Brown & French, 1979; Campione & Brown, 1984; Wozniak, 1975). This process involves an initial assessment of competence, followed by instruction on the target task(s). Children with high degrees of readiness (broad zones of proximal development) within that domain should benefit considerably from the intervention, whereas those with less readiness will not perform much better with this help than they did prior to it. As with other approaches, this measure of gain is presumed to possess greater predictive utility than the initial, unaided level of performance.

This framework has guided our work in both assessment and instruction. In this chapter, we will describe three sets of experiments that form part of an overall program of research with two major goals: (a) the development of diagnostic methods of assessing individual differences in students' readiness to perform in traditional academic domains; and (b) the use of the resulting information to guide the design of instructional programs that enhance the academic performance of students exhibiting relatively poor performance. In addition to the Vygotskian influence, they all involve aspects of dynamic assessment. Despite these similarities, the series also differ from each other in important ways. The differences arise because the studies are addressed to different issues within the present enterprise, including some that are primarily of theoretical interest and others that involve both theoretical and practical issues. Before proceeding to the specific studies, we will review some of the considerations that influenced our specific choices.

Background

To put the overall research effort into context, we will describe the issues that have attracted our interest over time and the considerations that led us to this particular approach. We have long been concerned with the diagnosis and remediation of weak students' academic problems. To do this, we need: (a) to identify the students likely to experience difficulties; (b) to analyze the academic domain in question in terms of a theoretical specification of the skills underlying successful performance; (c) to apply methods of assessing the individual's competence with those skills; and (d) to implement instructional methods for overcoming whatever deficiencies may be revealed through the assessment process.

As with many others, we have been less than optimistic about the role standard ability tests can play in this overall endeavor. In the next sections, we outline several reasons for this concern. The first involves a contrast between two different kinds of diagnostic procedures. This is followed by a more detailed analysis of the structure of standard tests. Having then described our reasons for adopting dynamic assessment methods, we distinguish two distinct uses we have made of the term.

Forms of Diagnosis

With regard to diagnosis, there are two levels at which the enterprise can be evaluated, one mainly aimed at identification and the second more concerned with prescription. In the former case, we might be concerned with identifying the students who are likely to experience difficulties, thus indicating the need for particular attention. A more valuable diagnosis would also be prescriptive; it would specify in detail the reasons for the problem, thus indicating both the need for, and direction of, remedial attempts. Although both identification and prescription are valuable, prescription enables us to work toward the second, instructional goal.

This distinction highlights the strengths and weaknesses of standard intelligence and ability tests. Under some circumstances, they do provide information that contributes to the identification goal, i.e., they can indicate students who are likely to experience problems; however, even this success has its limitations. Of more importance, standard tests have been much less successful at meeting the prescriptive goal. In the next section, we review some hypotheses about the sources of the specific limitations of those tests.

<u>Limitations of Standard Test Procedures</u>

One immediate question which arises is why there is a need for dynamic assessment approaches. Our goal is to link diagnosis and remediation; however, that goal is by no means novel. Standardized intelligence and ability tests were intended to identify individuals with academic problems and many were designed to provide "profiles" of ability that should allow a somewhat detailed analysis of the strengths and weaknesses of individuals. From such a picture, it should be possible to prescribe interventions tailored to the needs of particular students or groups of students. Such approaches, however, have not yielded much in the way of encouraging results (Brown & Campione, in press; Mann, 1979), and there are several reasons why this might be the case:

Product-based nature. Standard tests analyze the student's current level of performance but provide no direct evidence regarding the processes that may have operated or failed to operate to bring about that performance. As such, they provide at best a partial view of the testee's status. A nice statement of this point was made by Vygotsky, who noted that static test scores do not provide any information about

> those functions that have not yet matured but are in the process of maturation, functions that will mature tomorrow but are in the embryonic stage. These functions could be called the 'buds' or 'flowers', rather than the fruits of development. The actual developmental level characterizes mental development retrospectively, while the zone of proximal development characterizes mental development prospectively. (Vygotsky, 1978, pp. 86-87)

It is not that developers of standard tests are unconcerned with process--they do interpret the results in terms of sets of processes--but rather that their approach is to infer the processes underlying test performance from analyses of the structure of the test results themselves. Given their nature, standard tests rest heavily on the assumption that all testees have had comparable backgrounds and opportunities to acquire the information requested. This assumption is particularly suspect for students from minority or disadvantaged backgrounds. With such populations, abilities are quite likely to be underestimated. The result is that the identification goal of the evaluation is jeopardized.

Level of description. If we turn to the prescriptive aspects of assessment, there further problems with the process analyses involved in traditional tests. The "profiles" that result from such tests and that are used as the basis for description and diagnosis are couched in terms of very global aspects of performance (e.g., auditory sequencing) that are not easily theoretically relatable to interesting academic areas and tasks. Such diagnoses at best rest on somewhat vague abstractions from a particular psychological theory and cannot provide the kind of specific information needed to design instructional programs. For example, if auditory sequencing were diagnosed as the problem, it is not clear how best to intervene. Even if such skills can be developed, it is then left to the student to determine how and when these skills are to be used in academic contexts.

Degree of generality assumed. Finally, there is a related problem. The profiles that emerge are based on assumptions about the generality of the factors inferred from such tests. The abilities are presumed to be extremely general ones that operate in many, if not all, academic domains. While domain-general skills may well exist, it is also abundantly clear that there are important domain-specific capabilities that underlie successful performance in different domains, e.g., mathematics or reading. The tests available generally do not tap these skills in any meaningful way. While perhaps obvious, it seems reasonable to argue that if one is interested in assessing skill in the area of math, the assessment should be situated in the context of math problems. Again, the case of such processes as auditory sequencing is illuminating. The potential relevance of these processes to intervention programs rests on very strong theoretical assumptions about the nature of academic intelligence. The factors are presumed to be quite general, with the result that they affect performance in many situations. Improving auditory sequencing, then, would be expected to have widespread effects throughout the system. The analogy is with a muscle system in which practice on different skills strengthens the overall system and thus affords generalized improvement in performance.

Static nature of evaluation. Although not a necessary feature of standard tests, nonetheless the result of assessment is frequently taken as providing a relatively permanent characterization of the individual in question. The classifications that result, already presumed to reflect "general" academic ability, further tend to be regarded as fixed and unlikely to change over long periods of time. A measured IQ of 70, for example, is frequently assumed to reflect a relatively permanent characteristic of the student in all situations and under all circumstances.

Interpretations of Dynamic Assessment

As this volume is concerned primarily with issues regarding dynamic assessment, we feel it useful at this point to contrast two different ways in which we have used the term. The question is, what is dynamic about dynamic assessment? Although in both cases the important distinction is between static and dynamic properties of the assessment process, they differ in what, within the procedure, is regarded as dynamic--that being assessed, or the assessment itself. In the more traditional usage, the one we have already described, the interest is in assessing the efficiency of operation of the psychological processes involved in growth and change. The interest is not so much in evaluating an individual's current state of knowledge or skill as in estimating his or her readiness for change. The contrast is clearly with standard test procedures in which descriptions of individuals are couched in terms of what they currently know about some domain, or alternatively stated, between product- and process-based assessments of individual differences.

In the second case, we emphasize the dynamic nature of assessment itself—the notion that any assessment needs to be continuously re-evaluated as the student begins to acquire skill within some domain. Again, this is a feature of Vygotsky's treatment of the zone of proximal development. His argument is that instruction creates this zone; hence, with instruction, an individual's zone of proximal development changes, and it becomes necessary to continually update the diagnosis if instruction is to be appropriately directed. The assessment of an individual's zone, or readiness, is assumed meaningful for only brief periods, as one's readiness can itself change with practice and/or instruction. In this vein, we have also attempted to construct situations in which the assessment itself is dynamic rather than static, cases where the evaluator continually refines the diagnosis of the learners as they acquire competence. Thus, we use the phrase dynamic assessment to refer to: (a) assessment of process, or of the dynamics of change; and (b) to the need to continually change and refine the diagnosis of the individual learner, i.e., the dynamic, constantly changing nature of assessment itself.

An Alternative Approach

Our approach to dealing with the limitations of standard test methods involves several features. One is that assessment should evaluate as directly as possible the particular processes underlying successful performance. The second is that the assessment should ideally be situated within a specific domain, rather than being aimed at "general intellectual functioning." This in turn increases the likelihood that the processes can be specified in sufficient detail that instructional prescriptions can be designed. Finally, we make explicit the assumption that any diagnosis may have a very short half-life, and that re-diagnosis must be an integral part of any resultant intervention.

Having decided to concentrate as directly as possible on process is only a first step--it is still necessary to specify the process(es) to be evaluated, and then to determine how to situate that assessment.

Target processes. In our work thus far, we have concentrated on the role of quite global learning and transfer processes; the long-term goal is to be much more specific about the factors underlying individual differences in learning and transfer. In initial studies, we looked at the extent to which these global processes were related to overall academic ability. In more recent studies, we have concentrated on learning and transfer processes assessed within specific domains. In effect, we have assumed that estimates of individuals' learning potential and transfer efficiency within some domain provide measures of their readiness to perform in that domain.

This view emerged from a long series of studies with scholastically weak students, frequently labeled as learning disabled or mildly retarded. In that work, we concluded that in a variety of problem-solving situations, those students had difficulty learning new information (required complete and detailed instruction to do so) and were relatively unlikely to use that information flexibly in new problem situations (Brown, 1974, 1978; Brown & Campione, 1978, 1981, 1984; Campione & Brown, 1977, 1978, 1984; Campione, Brown & Ferrara, 1982).

Methods of assessment. Although that seems reasonable enough, at the time we began this program, the bulk of the available evidence did not support the position that assessments of learning ability or transfer flexibility would provide much helpful information about individual students (see Campione et al., 1982; Campione, Brown & Bryant, 1985). The question is how one might reconcile the disparate sets of findings. We have outlined our hypotheses in other sources (e.g., Campione & Brown, 1984), and will summarize them here. The major argument is that, in the studies generating negative findings, the estimates of learning and transfer efficiency were obtained: (a) in asocial learning situations; (b) involving only minimal feedback from the evaluator, most frequently simple feedback about the correctness of individual responses; and (c) situated in arbitrary domains. The metrics of learning and transfer were the amount of time and/or the number of trials needed to bring about learning.

As an example of this research approach, consider some comparative studies reported by Woodrow comparing the learning (Woodrow, 1917a) and transfer (Woodrow, 1917b) performance of groups of retarded and nonretarded children with mental ages of around ten years. The learning tasks he used involved a geometrical form sorting task in which the children were required to sort five forms into different boxes. They sorted 500 of these a day for 13 days, guided at best by feedback about the correctness of their individual placements. The main index of learning was the increase over time in the number of forms sorted. The transfer tasks consisted of two new sorting tasks (lengths of sticks and colored pegs) and two cancellation tasks (letters and geometric forms). Using these tasks, Woodrow found no differences whatsoever between the retarded and nonretarded groups in either learning or transfer performance. In these studies, learning and transfer were seen as passive, asocial, extremely general processes that could be tapped in any task domain. These conditions were typical of many studies failing to find evidence that learning and/or transfer processes represented important dimensions of individual differences (see Woodrow, 1946, for a review).

In contrast, the more recent studies, those yielding positive results, are characterized by a concern for structured intervention, often involving complex social interaction. The problems to be learned are set in non-arbitrary domains, i.e., ones where there are rules for the students to learn and where it is possible to come to understand why certain responses are appropriate in given situations and not in others. This understanding then serves as the basis for subsequent use of the newly acquired information, i.e., principled transfer is possible (Brown & French, 1979; Campione & Brown, 1984). The metric of learning or transfer efficiency is the amount of help needed for a student to acquire a rule or procedure.

Given this analysis, we assumed that if we wished to assess individual differences in learning and transfer that would be of diagnostic significance, we would have to match these latter conditions. The learning should be guided by the adult tester and should involve the acquisition of rules or principles whose application in novel contexts we could subsequently observe. These ideas clearly meshed nicely with those of Vygotsky (1978), and our procedures have ended up being quite similar to those employed by neo-Vygotskians in the Soviet Union.

The studies we have conducted follow the same general format. They begin with an evaluation of children's initial competence. Following this, they are placed in a mini-learning environment where an adult (or a computer) works collaboratively with them until they are able to solve sets of problems independently. If they are unable to solve a particular problem, they are given a series of hints to help them. The initial hints are very general ones, and succeeding ones become progressively more specific and more concrete, with the last "hint" actually providing a detailed blueprint for generating the correct answer. This titration procedure allows us to estimate the minimum amount of help needed by a given child to solve each problem. The metric of learning efficiency is the number of hints required for the attainment of the learning criterion (typically two successive problems solved with no help). Note that the metric here differs from that used by several others interested in dynamic assessment, including Vygotsky, in that it is not how much improvement one can bring about through intervention, but rather how much aid is needed to bring about a specified amount of learning.

Exactly the same hinting procedure is used on the transfer problems, generating the analogous metric. Note that the index of transfer propensity is thus a dynamic, rather than static, one (Brown, Bransford, Ferrara & Campione, 1983). That is, we do not measure how many and what types of transfer items individuals can solve on an unaided test (a static measure), rather, we are concerned with how facile they are in coming to deal with related portions of the overall problem space (a dynamic measure)--specifically, how many hints they require to solve the various types of transfer problems. Following these instructional sessions, a post-test is given, and the gain brought about by the instruction determined.

To summarize, we decided to situate our assessment of learning and transfer efficiency in social interactional contexts in which the evaluator would be engaged in the task of teaching the children how to solve sets of problems; the measures of learning and transfer could then be based on students' responses to that instruction.

There is one further point to emphasize. The hints employed were based on a detailed task analysis and were designed such that each one would provide more specific information than the previous one(s). These hints were given in a fixed sequence and were, with one exception, independent of the individual child's responses (the exception was that if the child had already generated the information provided by an early hint, that hint was omitted, and the experimenter gave the next hint in the sequence). The procedure was then _task-_, rather than _child-oriented_. This was done because we aimed to produce quantitative data with good psychometric properties; the amount of help indices are likely to have such properties only if the test administration is standardized as much as possible.

The trade-off is with more clinical procedures in which assessors vary their questions, or prompts, with different children as those children show different approaches to the problems at hand. Such approaches may well provide richer information about the skills and aptitudes of individual children; however, they are less likely to produce strong quantitative data. As we will describe later, we have attempted in some of our more recent studies to modify our procedures in ways that allow us to combine the strengths of the different approaches.

Task domains. The next decision involved the selection of a domain in which to embed the teaching. Given an interest in transfer propensity, it is necessary to choose a domain in which rules or principles can be learned and applied to novel types of problems. As we were also interested in academic skills, we also wished to choose a domain that was known to be related to school performance. In our initial studies investigating the diagnostic utility of measures of learning and transfer, we worked with inductive reasoning problem spaces, variants of progressive matrices problems and series completion problems, as performance on those tasks is known to be related to scholastic success. Further, enough was known about the structure of those tasks that it was possible to design a theoretically-based teaching, or hinting, sequence. In our work on instruction emphasizing the dynamic nature of the assessment process itself, we concentrated on studies of reading and listening comprehension, skills of considerable academic importance.

Specific Studies

We will summarize the results of three sets of studies. The first two involve the theoretical and diagnostic status of the learning and transfer measures obtained in our adaptation of the zone of proximal development testing procedures. These deal with issues of <u>concurrent validity</u> and <u>predictive validity</u>. In the first case, we selected students of varying academic ability and assessed their performance as they learned how to solve inductive reasoning problems. Performance on these problems, featured on most ability tests, consistently distinguishes academically successful from less successful students. Our expectation, then, was that learning and transfer indices, obtained in these domains, would be related to assessed ability.

In the case of predictive validity, we wished to go one step farther and evaluate the extent to which the dynamic measures would provide diagnostic information beyond that afforded by static ability tests. While ability test performance was expected to be related to learning/transfer efficiency in these inductive reasoning domains, we also expected that the dynamic measures would provide more information about the future performance of subjects within those domains than would the static tests.

In these studies, instruction is provided, and response to instruction is used as a metric of individual students' readiness to deal with the domain under study. The concern is with how much instruction is needed to bring about a given level of performance.

The third series is more "purely" instructional. The goal here is to maximize the performance of individual students in important academic domains. One key element of the instructional program is the need for continually updating the diagnosis of students' current skill levels, rather than using the initial estimate as a long-term index.

Studies of Concurrent Validity

In these studies, we were interested in the extent to which measures of learning and transfer efficiency, obtained within the context of prototypic ability test items, specifically inductive reasoning tasks, would be related to general ability levels. There are two issues involved: (a) regarding diagnosis, do either or both measures distinguish lower ability students from those of higher ability--an identification issue? and (b) regarding theory, can part of the variance in individual differences in this domain be attributed to learning and/or transfer dynamics--a qualitative issue.

Campione, Brown, Ferrara, Jones, and Steinberg (in press). In this study, we used a variant of the Raven Progressive Matrices task. At the outset, subjects were given a pre-test involving the kinds of problems that were to be used in the diagnostic/instructional sessions. Each problem involved a 3 x 3 matrix wit the lower-right entry left blank; the subject's task was to select, from a set of six, the pattern that best completed the matrix. The subjects consisted of groups of retarded (mean IQ = 72) and nonretarded (mean IQ = 118) children matched for a mental age of approximately 10.5 years and for performance on the pre-test.

During the instructional sessions, students worked at a computer terminal. In the initial phase, they learned to solve problems involving three rules: rotation, imposition, and subtraction. Examples are shown in Figure 1. During this, the learning portion of the study, the problems were presented in a

---

Insert Figure 1 about here

---

blocked format. Each student learned the rotation problems to a criterion, then the imposition problems, and finally the subtraction problems (an easy-to-hard sequence). In the next (maintenance) session, novel exemplars of the same type were presented, but now in a random order. The ensuing (transfer) session included these same problem types interspersed with a set of transfer problems; these required the use of combinations of the original rules (rotation + imposition; rotation + subtraction). Examples are shown in Figure 2.

Insert Figure 2 about here

In contrast to the pre-test procedure (and the standard procedure used with the Raven), their task here was to <u>generate</u> the pattern needed to complete the matrix by issuing a set of pre-taught commands using a touch-sensitive screen. Graduated and animated hints were provided via the computer as needed, with an adult reading the hints to the child if necessary and providing general encouragement. The hints were presented in a pre-set sequence, proceeding from very general hints offering relatively little specific information about the form of the solution to very specific hints, which eventually provided a detailed blueprint from which the child could generate the correct answer. The numbers of hints taken to reach the learning criterion and to solve the maintenance and transfer problems were the metrics of learning and transfer efficiency. A sample hint sequence (for Rotation problems) is shown in Figure 3.

Insert Figure 3 about here

No differences were obtained during the learning phase of the study, possibly due to the matching procedures that equated the groups for both mental age and entering competence. However, group differences were apparent during both the maintenance and transfer phases of the study. Further, those differences tended to increase as the similarity of the training and test contexts decreased. The greater the need for flexibility in applying the learned rules, the larger were the differences between retarded and nonretarded children. Thus, in this study, transfer, but not learning, performance did distinguish the different ability groups.

Ferrara, Brown and Campione (in press). This study included a different ability range and a developmental factor, contrasting third and fifth grade children of average and above average ability. A second inductive reasoning task, letter series completions, was used. One major difference between the series completion and matrices tasks was that in the series completion case, a more detailed examination of "transfer distance" was included. The idea, supported by the results of the Campione et al. study, was that individual or group differences would be more apparent as the transfer distance, or the difference between the learning and transfer situations, increased. While there have been some suggestions that, for example, retarded children may show "near transfer," they are quite unlikely to show "far transfer" (e.g., Brown, 1978; Campione & Brown, 1977). A nice general statement of this notion was provided by Gagne (1970) in the course of describing lateral transfer:

> In the case of this kind of transfer, the question of how much appears to be a matter of how broadly the individual can generalize what he has learned to a new situation. Presumably, there are limits to the breadth of generalization, which vary with different individuals. One could perhaps think of a whole range of situations of potential applicability of (some learned rules) that display decreasing degrees of similarity to the situation in which the rule had

> originally been learned. At some point along this dimension of breadth of generalization, a given individual will fail to transfer his previously learned knowledge. Another individual, however, may be able to exhibit transfer more broadly to a wider variety of differing situations. (p. 336)

While the idea is an attractive one, there are few relevant data available, one problem being that there has frequently been no objective way of determining transfer distance in the domains that have been investigated. The series completion task was one that lent itself nicely to this task. Specifically, "transfer distance" can be defined in terms of the number of transformations distinguishing the learning problems from the various transfer items. Examples of the learning, maintenance, near transfer, far transfer, and very far transfer items are shown in Table 1.

Insert Table 1 about here

The child's task is to fill in the blanks with letters that continue the pattern that is determined by a certain periodicity and by certain alphabetic relations (Next, the appearance of letters in alphabetical sequence; Identity, the repetition of letters; and Backward-next, the appearance of letters in reverse alphabetical sequence). On the learning items, the children learned to deal with the Next and Identity relations, and with periodicities of two and four. Maintenance items involve no transformations, but are simply novel exemplars of the same problem types learned originally. Near transfer items involve the same principles (relations and periodicities) learned originally but in different combinations. Far transfer items involve the application of a novel periodicity (three) or relation (Backward-next). And Very far transfer items involve the use of novel principles in a novel context.

There was an overall effect of ability during the learning phase; high ability children needed fewer hints to learn the initial problems than did the average ability group. The transfer data, however, were of more interest. The major finding, as expected, was that group differences increased as transfer distance increased. These results are shown in Figure 4 where it can be seen that virtually no aid was required on the maintenance items and very little on the near transfer items; there were no instances of group differences. However, on the

---

Insert Figure 4 about here

---

far and very far transfer items, group differences were highly reliable. The results of a series of correlational analyses revealed the same pattern. Correlations between IQ scores and number of hints taken were non-significant for maintenance and near transfer, but reliable when far and very far transfer performance was considered.

Campione and Ferrara (in preparation). The next study involved a comparison of retarded and nonretarded children on the series completion task. The results were quite consistent with those of the first two studies. Group differences emerged on both learning and transfer sessions, with the differences being larger during transfer. As in the previous studies, the nonretarded children performed extremely well on the maintenance series, requiring virtually no help to solve those problems; however, the retarded students did need experimenter-provided hints to deal with those problems. On far and very far transfer tests, the differences between the groups again increased reliably.

Overall, the results of these three studies establish the concurrent validity of the learning and transfer measures. Groups of children of contrasting ability do differ in terms of their learning, and particularly transfer, performance. Less able children tend to need more help to solve sets of original learning problems, and then continue to be at a disadvantage when they are required to make flexible use of the rules or principles they have been taught. The greater the amount of flexibility required, the larger is the difference between the groups.

The notion of transfer distance does appear to be an important one in terms of diagnosis of group differences. The farther the distance, the larger the magnitude of any difference. Further, groups of different ability vary in how "far" they can transfer before they begin to run into difficulties. In both Studies 1 and 3, where retarded children were involved, those students began to require help even on the maintenance series. Having learned to solve particular sets of problems, they still run into difficulties when they are asked to solve problems of the exact same type later in a different context. Children of average and above average ability, in contrast, handle maintenance (and even near transfer) items extremely well. It is only when far transfer problems are given that they begin to need significant amounts of help; and only when these far transfer problems appear do the average and above average ability groups begin to differ.

Studies of Predictive Validity

In the next set of studies, we (Bryant, 1982; Bryant, Brown & Campione, 1983): (a) attempted to extend the previous results to younger children, five-year-olds; and (b) addressed the issue of whether the learning and transfer measures do provide additional diagnostic information about individual subjects beyond their standard ability scores. The ideal way to evaluate predictive validity would be to have measures of initial competence, along with measures of general ability and learning and transfer scores. Then, at some later point in time, we could re-assess the students' ability. The question would then be which score(s) best predict later performance--initial competence, general ability, or the learning and transfer indices. In these studies, our general procedure was to give subjects a pretest, learning and transfer sessions, and then a final posttest. The transfer sessions included maintenance, near transfer, and far transfer items, again defined in terms of the number of transformations distinguishing the transfer probes from the learning items. The pretest included both evaluations of general ability (subscales from the WPPSI to generate an overall IQ estimate and the Raven Coloured Progressive Matrices) and a task-specific pretest. In the latter, baseline levels of performance on the items to be included in the learning and transfer sessions were obtained. The posttest was a re-administration of this pretest, and our major interest was in the gain that resulted as a consequence of the instruction afforded in the learning and transfer sessions.

Two separate studies were conducted, one involving a simplified version of the matrices task (Bryant, 1982) and the second a simplified version of the series completion task. The major results are shown in Tables 2 and 3, which portray a series of multiple regression analyses. The first thing to note is that there are significant relations between the ability scores and the learning and transfer metrics, thus replicating the results of the previous studies. Children of higher ability tend to require fewer hints to solve the original sets of problems, and further require fewer hints to deal with the transfer problems.

Insert Tables 2 and 3 about here

Of more interest are the results of the analyses of the gain scores. In these analyses, the effects of the estimated IQ score and the Raven score were extracted first. In both studies, they did allow a reasonable prediction of the gain score, accounting for around 60% of the variance in that score. Even after the effects of the ability scores were extracted, however, the learning and transfer scores still accounted for significant additional portions of the variance in gain scores; thus, taking the learning and transfer scores into account did provide further diagnostic information about individual children. In the matrices task (Table 2), the learning score accounted for an additional 22% of the variance, the transfer score for a still further 17%. In the series completion task, the learning score did not result in an increase in predictability of gain scores, but the transfer score did account for an additional 22% of the gain variance beyond the ability scores. Alternatively, if one looks at the simple correlations, the learning and transfer scores are better predictors of gain score than either of the static ability measures. Finally, within the set of dynamic measures, the tendency is for the transfer measures to be more strongly associated with gain scores than the learning index. This is consistent with the findings from the earlier series where ability group differences were larger on transfer than during learning.

The studies reviewed thus far establish that the dynamic assessment measures do provide diagnostic information about children and can play a role in the <u>identification</u> component of diagnosis. But what about the <u>prescriptive</u> component, i.e., do they suggest any particular sources of problems to which instructional programs might be geared? The answer is, "Yes, to some extent." Throughout the series of studies, the largest and most consistent effects have had to do with aspects of transfer propensity. Transfer measures were most strongly related to ability measures--it is transfer that best discriminates various ability groups. And transfer flexibility is also the best predictor of gain scores. Our best overall description of differences between more and less successful students would then be in terms of the processes underlying the judicious application of acquired skills to the solution of novel problems. The suggestion is that any programs that are designed for use with academically weak students must deal with the transfer issue. It is not sufficient to plan instruction in such a way that rules and principles are learned to some criterion; it is also necessary to attempt to provide these tools in a context that stimulates students' ability to use them with some flexibility.

Although this is a general suggestion, and one that has been around for decades, i.e., teachers are frequently told to "teach for transfer" even if not taught how to do so, what makes it more than a platitude is the fact that some of the instructional principles that are effective at inducing transfer have been identified (e.g., Brown & Campione, 1978, 1981; Palincsar & Brown, 1984) and shown to be effective. These include training in multiple settings, attention to the metacognitive environment of instruction--making the student aware of the skills being taught and of the need to actively monitor and regulate them, and the range of applicability of those skills--teaching the skills in the actual context in which they are to be used, rather than as isolated skills, etc. Further, many programs designed for weak students intentionally <u>do not</u> include such components; the idea is that for such students instruction should concentrate on making sure they "know the basic facts," a form of mastery learning that leads to a concentration on drill aimed at perfecting individual skills, quite the opposite of the conclusions we have reached. To buttress this argument, in Section D, we review a program of research that has embodies these features and that has produced impressive results. That work has included the general suggestions mentioned here along with more specific suggestions that followed, once the particular domain in question had been specified. Before turning to the instructional work, however, we would like to indicate the ways in which we are attempting to improve the basic assessment procedures.

Current and Future Issues Regarding Assessment

The initial results obtained in our adaptation of Vygotsky's zone of proximal development approach to assessment have been encouraging. Over a series of studies involving different tasks and subjects of widely varying ages and abilities, the learning and transfer metrics have consistently provided useful information about students. They are related to ability measures (which are themselves predictive of academic success), but also provide additional information not captured by those tests. They have also consistently led to an emphasis on transfer processes as sources of individual differences, and hence to suggestions about the design of intervention programs aimed at weak students. Our ongoing studies include attempts to improve on the diagnostic properties of the dynamic measures, in terms of both the identification and prescriptive goals. We are also extending the procedures to richer, more academically relevant tasks for both practical and theoretical reasons.

The role of personality factors. One line of research attempts to improve on the predictive power of dynamic measures by adding information about individuals' attribution styles. The assessment is carried out in a social, interactional system where an expert and a novice work together to solve sets of problems. In this situation, students ask for, and are given, help as needed. It is unlikely that their responses in such a social situation are determined by purely "cognitive" factors. On a general level, there appeared to be clear differences between the ways young children (five-year-olds in Bryant's work) and the elderly (French, 1979) responded to the hinting procedures. The elderly appeared threatened by the need for hints and interpreted them as indicating that they were failing on the task; in contrast, the young children appeared more willing to accept the help and still feel that they had solved the problems themselves. There is also some evidence that different children interpret the input in different ways, some seeking it frequently and others doing everything they can to avoid asking for help so they can in fact be allowed to solve the problems themselves. This, along with some of their spontaneous verbal comments, indicates that children adopt either learning or performance goals (Dweck & Elliot, 1983; Dweck & Bempechat, 1984) in the task that lead them to react differently to the need for aid. If we could assess those orientations, we should be in a better position to evaluate their performance during the assessment sessions. To collect some relevant data, we have re-done the Campione et al. (in press) matrices study with a large group of fourth graders. We have also administered these students a pair of social comparison questionnaires--Crandall, Katkovsky and Crandall's (1965) Intellectual Achievement Responsibility Scale and Harter's (1983) Perceived Competence Scale: Revised Version. The hypothesis is that by taking into account individual's attributions and orientations, more accurate predictions about future performance can be obtained.

Qualitative analyses. We are also attempting to generate richer descriptions of individual students and more detailed and prescriptive pictures of the differences between successful and unsuccessful students. We would like to be able to get rich qualitative descriptions of students' approaches to the problems while maintaining the standardized format that has produced useful quantiative data, i.e., to merge the psychometric and clinical approaches.

One approach that is particularly promising involves having students "talk-aloud" about their approaches as they work on the problems. The initial attempt here has shown that fourth graders can handle this requirement quite well, and it appears that adding this component does not materially change the ways in which they approach the problems. (It is also the case that some five-year-olds provide spontaneous talk-alouds during the testing sessions; the social interactional nature of the assessment process seems to support this nicely. When these talk-alouds do occur, they are quite informative about individual children's approaches and supplement the quantitative data in interesting ways.) Although these data are not fully analyzed, these talk-alouds do appear to provide useful information. For example, successful students tend to spend a considerable amount of time "planning" their moves; they talk about what the answer should look like before they begin to construct their own answer or bef e the consider the alternatives from which they have to select. Less successful subjects tend to begin their construction activities without fully analyzing the problem; they tend to misclassify problems and proceed in an unsystematic fashion to construct an answer. There are also differences in the ways in which successful and unsuccessful students recover from errors, or from false starts that do not lead to problem solution. We are confident that this information could be used to design more powerful and individually tailored programs of instruction.

Extension to academic domains. Finally, having shown that our procedures do work, we are extending them to richer and academically more interesting domains--initially early mathematics. There are several reasons for doing this.

First, pragmatically, this domain is one of clear educational significance, and given that the procedures we have used require a large amount of effort to develop, it makes sense to situate that work in such an area. Assessments of individual readiness are of more immediate interest if obtained in math than in inductive reasoning domains; and the leap from diagnosis to suggestions for the design of instructional programs is shorter in the case of math.

Second, of more theoretical interest, there have been a number of detailed analyses of the structure and development of mathematical knowledge. This work makes it possible to obtain a reasonably thorough picture of students' mathematical knowledge before they enter the assessment situation. This is important because it is only if we can assess the quality of an individual's knowledge in some area that we can clearly evaluate the differential contributions of content knowledge and learning/transfer dynamics to the assessment process. For example, some might argue that the differences in learning and transfer efficiency which are uncovered in our studies are actually no more than manifestations of individual differences in content knowledge. Unless we have a good measure of that knowledge, it is difficult to refute that claim. In some current work, Ferrara, as part of her dissertation research, is working on the development of a test of earl math knowledge. She is also designing hinting procedures that can be used with simple addition and subtraction problems. With these assessments of knowledge and learning/transfer efficiency in hand, it will be possible to assess the predictive properties of the dynamic measures when students are equated for their entering knowledge.

This leads to the last point. We are interested in devising measures that can predict students' future trajectories. The success of such an enterprise can best be evaluated in an area where there is room for a large amount of improvement; inductive reasoning problems of the type we began this research with do not serve this purpose well. In contrast, mathematics, even early mathematics, is an area where there is considerable room for improvement; as such, we can track the progress of students over long periods of time while they are acquiring increasingly sophisticated sets of skills. In this way, stronger tests about the utility of dynamic assessment procedures can be arranged.

Instructional Design: The Dynamic Nature of Assessment

In this section, we describe the highlights of a program of research that has concentrated more directly on instruction, and instruction in a particular academic domain. The concentration on a specific domain makes it somewhat easier to specify in more detail the skills distinguishing strong from weak students. The goal was to improve the reading and listening comprehension skills of students experiencing particular difficulties with that task. In the main studies (Brown & Palincsar, 1982, in press; Palincsar & Brown, 1984), the students were seventh grade students of relatively low overall ability (IQs ranged from 60-100, with a mean of around 80) whose reading comprehension scores lagged one to four years behind those of their age- and grade-mates.

The general design of the instructional sessions was based on the same Vygotskian principles that guided the development of the assessment procedures. We sought to mirror in the teaching situation the gradual transfer of control of cognitive skills that Vygotsky described. The teacher and students would begin by working together, with the teacher initially doing most of the work. As the children began to acquire the target skills, they were encouraged to take on more and more responsibility until they were eventually able to employ the skills when working independently.

More specific features of the instruction was based on a considerable amount of prior research indicating that one major difference between skilled and unskilled comprehenders lay in the kinds of active comprehension strategies (both comprehension-fostering and comprehension-monitoring--see Palincsar & Brown, 1984, for additional description) they brought to the task of reading for meaning. Specifically, good readers, in the course of studying a text, tend to: (a) stop and summarize what they have read periodically; (b) formulate questions that capture the main idea of what they have just read; (c) attempt to clarify any inconsistencies that appear; and (d) predict what the author will go on to say. The instructional program that Palincsar and Brown developed was designed to teach these four strategies. Our interest here is with only a portion of the overall program, the way in which assessment of student capabilities is integrated into the overall framework.

If one were to consider the students with comprehension problems and diagnose their competence in the use of the four activities just listed, it would turn out that all would essentially fail the test. There is little evidence of their using these activities without explicit instruction. Further, when asked to engage in the activities, they do so very poorly. We have then a reasonable diagnosis about the sources of their problems. The question is, what do we do about it? How strongly should that diagnosis affect instruction? And how long should that diagnosis be retained? The procedures that Palincsar and Brown developed, termed reciprocal teaching, provide some insights into these issues.

We do not have the space here to provide the details of their approach. The main point for our purposes is that in the teaching sessions, the teachers engaged in constant on-line diagnosis and re-diagnosis of each student's current level of skill. This was possible because the teaching method forced each student to produce the key activities overtly. When the group was engaged in reading a text, the teacher and students took turns leading a dialogue about the text segment they had just read. The leader of that dialogue was required to summarize what had just been said and formulate a question about the main point of the section. When appropriate, they were also told that they should seek to clarify any inconsistencies or confusions that arose, and to predict what might happen next.

Several features of this interaction are important; they were included to maximize the likelihood that transfer of the target strategies to an array of academic tasks would result.

First, note that the students engage in the target activities in the context of actually reading and understanding texts. It is also made clear to them what those activities are, why they are useful, and where they can be applied. Further, as the students carry out the activities, the teacher is able to see how well they are executed and diagnose what individuals' current problems, if any, with the particular skills are. In this way, feedback can be provided to each student tailored to particular needs at the moment.

Over time, as student competence increases, the teacher's diagnosis changes, and different types of feedback are provided requiring more advanced responding from the student. In this way, the student is gradually led to master the various activities, until eventually an acceptable level of skill is reached. The teacher begins by doing a large part of the work for the student, but as the diagnosis changes, progressively more work from the child is required until the teacher can eventually fade out, leaving the student to perform unaided. Our point here is simply that the initial diagnosis (the students do not engage in these activities) needs to be constantly updated, so that the teacher can respond appropriately to the students' needs at any point in time, and thus provide the kind of input necessary to move them one step further toward independent competence.

To see how this works in practice, we can consider a classroom teacher interacting with two remedial seventh graders--

her interactions are quite different in the case of Charles, who makes a very weak beginning, and Sara, who has a clear (but inadequate) idea of how to ask questions concerning texts. (Charles, IQ = 70, Reading Comprehension = third grade; Sara, IQ = 84, Reading Comprehension = fourth grade.)

Charles. The group is reading a passage about American snakes. Charles has a great deal of difficulty taking his turn leading the dialogue, primarily because he doesn't know how to formulate an appropriate question (see Table 4). "What is found in the Southeastern snake, also the copperhead, rattlesnakes,

----

Insert Table 4 about here

----

vipers--they have--I'm not doing this right." The teacher responds to his difficulty and tells him the main idea. "Do you want to ask something about the pit vipers?" When he still fails to ask an adequate question, she prompts, "Ask a good question about the pit vipers that starts with the word why." When he still cannot manage it, she models, "Why do they call the snakes pit vipers?" After two tries, he copies the teacher's question and she provides praise and encouragement. Even imitating a fully formed question is difficult for Charles initially.

Four days later Charles is still having difficulty asking questions on a passage about spiders. The teacher models one for him, but this time she waits for him to find the main idea himself and attempt to make up a question. "How do spinner's mate spend most of his time sitting?" The teacher responds, "You're very close. The question would be, 'How does spinner's mate spend most of his time?' Now you ask it." And he does.

Seven days into the procedure, Charles can make up questions with a little help pinpointing main ideas and by the eleventh day he takes his turn as teacher with two questions, "What is the most interesting of the insect eating plants," and "Where do the plants live at?" After fifteen days he produces acceptable questions each time it is his turn to lead the dialogue. Charles - "Why do scientists come to the South Pole to study?" Teacher - "Excellent question--that's what the paragraph is all about!"

Sara. In contrast to Charles, another student in the group, Sara (see Table 5) has a clear idea of what kinds of questions occur in schools--"fill in the blanks." The teacher, preoccupied with Charles, tolerates such questions until the second day and then attempts to take Sara beyond this level. Sara - "Snakes'

Insert Table 5 about here.

---

backbones can have as many as 300 vertebrates - almost _____ times as many as humans?" Teacher - "Not a bad beginning, but I would consider that a question about a detail. Try to avoid 'fill in the blanks' questions. See if next time you can find a main idea question and begin your question with a question word - how, why, when . . ."

On the third day, Sara comes up with a main idea question, but this time she selects a line in the text, "several varieties of snakes live all their lives in the sea," and turns it into a question, "Can snakes live their whole lives in seas?" The teacher again increases her demand and asks, "See if you can ask a question using your own words." For the remainder of the sessions, Sara composes questions in her own words becoming more and more like the model teacher in her turn.

The teacher's responses to Charles and Sara are different, and this variation appears to dovetail well with their entering skill levels and rates of improvement. As the teacher diagnoses their growing levels of competence, she asks more and more of them until they eventually generate good questions with no teacher guidance. Notice that they are never asked to make a large jump, never asked to move quickly to unaided performance. Rather, they are gradually guided to that level, something that can occur only if the teacher continues to update her assessment of their evolving capabilities.

This program has produced impressive results in a number of replications ranging from experimental studies involving small groups through larger-scale studies involving classroom instruction conducted by teachers with their regular, and frequently large, reading groups. We will highlight some of the major gains here (see Palincsar & Brown, 1984, for more detail). (a) Throughout the period during which instruction was provided, students took daily tests on their ability to read a science or social studies passage and then answer from memory ten comprhension questions. Instructed students' performance on these tests begins at around 30-40% correct and improves steadily until they are consistently scoring 80% correct. (b) There is also evidence that their newfound skills are being transferred to classroom activities. For example, in one study, all seventh graders in the school (approximately 140) took regular exams, consisting of reading passages and answering comprehension questions, in their science and social studies classes. At the beginning of the intervention, the students in the reciprocal teaching groups scored at around the 15th percentile; by the end they had moved up to above the 50th percentile. (c) They showed evidence of transferring some of the trained skills to laboratory-based tests. There were significant increases in their ability to detect text inconsistencies, generate questions probing the main idea of the passages they read, and write summaries of portions of assigned texts. And (d) their standardized reading comprehension scores increased significantly--by an average of just over two years.

Although none of the students showed evidence of using the target activities spontaneously at the outset of the studies, and some had extreme difficulty producing them when initially instructed to do so, the teacher was able to monitor the improvement that did occur and provide the kind of practice and feedback needed to continue that improvement. As a result, the students did learn to use the skills independently and flexibly, leading to worthwhile improvements in their ability to read and understand texts.

Summary

We have reviewed several lines of research incorporating features of dynamic assessment. In that research we have used dynamic assessment to refer to two distinct sets of activities, one emphasizing the view that assessment attempts should be aimed as directly as possible at the processes underlying successful performance on academic tasks, and the second that the assessment itself should be continuously updated. Our studies conducted thus far have shown that the measures of learning and transfer efficiency that we generate in our adaptation of Vygotsky's zone of proximal development testing procedures do possess both concurrent and predictive validity. They have also indicated that the best predictors of the extent to which individuals are likely to profit from instruction are their initial responses to instruction and, even more sensitive, the extent to which they can transfer their newly learned skills to novel situations.

In the context of instruction, we have argued that whereas early diagnoses can provide important information about the kinds of educational programs needed with weak students, those diagnoses need to be continuously updated if they are to contribute meaningfully to instructional goals. Diagnosis should not be used to pigeonhole students, but rather to provide information indicating how instruction should change over time.

References

Brown, A. L. (1974). The role of strategic behavior in retardate memory. In N. R. Ellis (Ed.), International review of research in mental retardation (Vol. 7, pp. 55-111). New York: Academic Press.

Brown, A. L. (1978). Knowing when, where, and how to remember: A problem of metacognition. In R. Glaser (Ed.), Advances in instructional psychology (Vol. 1, pp. 77-165). Hillsdale, NJ: Erlbaum.

Brown, A. L., Bransford, J. D., Ferrara, R. A., & Campione, J. C. (1983). Learning, remembering, and understanding. In J. H. Flavell & E. M. Markman (Eds.), Handbook of child psychology (4th ed.). Cognitive development (Vol. 3, pp. 515-529). New York: Wiley.

Brown, A. L., & Campione, J. C. (1978). Permissible inferences from the outcome of training studies in cognitive development research. Quarterly Newsletter of the Institute for Comparative Human Development, 2, 46-53.

Brown, A. L., & Campione, J. C. (1981). Inducing flexible thinking: A problem of access. In M. Friedman, J. P. Das, & N. O'Connor (Eds.), Intelligence and learning (pp. 515-529). New York: Plenum Press.

Brown, A. L., & Campione, J. C. (1984). Three faces of transfer: Implications for early competence, individual differences, and instruction. In M. Lamb, A. Brown, & B. Rogoff (Eds.), Advances in developmental psychology (Vol. 3, pp. 143-192). Hillsdale, NJ: Erlbaum.

Brown, A. L., & Campione, J. C. (in press). Cognitive science and learning disabilities. American Psychologist.

Brown, A. L., & Ferrara, R. A. (1985). Diagnosing zones of proximal development. In J. Wertsch (Ed.), Culture, communication, and cognition: Vygotskian perspectives (pp. 273-305). Cambridge, MA: Cambridge University Press.

Brown, A. L., & French, L. A. (1979). The zone of potential development: Implications for intelligence testing in the year 2000. Intelligence, 3, 253-271.

Brown, A. L., & Palincsar, A. S. (1982). Inducing strategic learning from texts by means of informed, self-control training. Topics in Learning and Learning Disabilities, 2(1), 1-17.

Brown, A. L., & Palincsar, A. S. (in press). Reciprocal teaching of comprehension strategies: A natural history of one program for enhancing learning. In J. Borkowski & J. D. Day. (Eds.), Intelligence and cognition in special children: Comparative studies of giftedness, mental retardation, and learning disabilities. New York: Ablex.

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.

Bryant, N. R., Brown, A. L., & Campione, J. C. (1983, April). Preschool children's learning and transfer of matrices problems: Potential for improvement. Paper presented at the Society for Research in Child Development meeting, Detroit.

Budoff, M. (1974). Learning potential and educability among the educable mentally retarded (Final Report Project No. 312312). Cambridge, MA: Research Institute for Educational Problems, Cambridge Mental Health Association.

Campione, J. C., & Brown, A. L. (1977). Memory and metamemory development in educable retarded children. In R. V. Kail, Jr., & J. W. Hagen (Eds.), Perspectives on the development of memory and cognition (pp. 367-406). Hillsdale, NJ: Erlbaum.

Campione, J. C., & Brown, A. L. (1978). Toward a theory of intelligence: Contributions from research with retarded children. Intelligence, 2, 279-304.

Campione, J. C., & Brown, A. L. (1984). Learning ability and transfer propensity as sources of individual differences in intelligence. In P. H. Brooks, R. D. Sperber, & C. McCauley (Eds.), Learning and cognition in the mentally retarded (pp. 265-294). Baltimore: University Park Press.

Campione, J. C., Brown, A. L., & Bryant, N. R. (1985). Individual differences in learning and memory. In R. J. Sternberg (Ed.), Human abilities: An information processing approach (pp. 103-126). New York: W. H. Freeman & Company.

Campione, J. C., Brown, A. L., & Ferrara, R. A. (1982). Mental retardation and intelligence. In R. J. Sternberg (Ed.), Handbook of human intelligence (pp. 392-490). Cambridge, MA: Cambridge University Press.

Campione, J. C., Brown, A. L., Ferrara, R. A., Jones, R. S., & Steinberg, E. (in press). Differences between retarded and non-retarded children in transfer following equivalent learning performance: Breakdowns in flexible use of information. Intelligence.

Carlson, J. S., & Wiedl, K. H. (1978). Use of testing-the-limits procedures in the assessment of intellectual capabilities in children with learning difficulties. American Journal of Mental Deficiency, 82, 559-564.

Carlson, J. S., & Wiedl, K. H. (1979). Toward a differential testing approach: Testing-the-limits employing the Raven matrices. Intelligence, 3, 323-344.

Crandall, V. C., Katkovsky, W., & Crandall, V. J. (1965). Children's beliefs in their own control of reinforcements in intellectual-academic achievement situations. Child Development, 36, 91-109.

Dweck, C. S., & Bempechat, J. (1984). Children's theories of intelligence: Consequences for learning. In S. G. Paris, G. M. Olson, & H. W. Stevenson (Eds.), Learning and motivation in the classroom. Hillsdale, NJ: Erlbaum.

Dweck, C. S., & Elliot, E. S. (1983). Achievement motivation. In E. M. Hetherington (Ed.), Handbook of child psychology (4th Ed.): Socialization, personality, and social development (Vol. 4, pp. 642-692). New York: Wiley & Sons.

Ferrara, R. A., Brown, A. L., & Campione, J. C. (in press). Children's learning and transfer of inductive reasoning rules: Studies of proximal development. Child Development.

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

French, L. A. (1979). Cognitive consequences of education: Transfer of training in the elderly. Unpublished doctoral dissertation, University of Illinois.

Gagne, R. M. (1970). The conditions of learning (2nd ed.). New York: Holt, Rinehart, & Winston.

Harter, S. (1983). Supplementary description of the self-perception profile for children: Revision of the perceived competence scale for children. Unpublished manuscript, University of Denver, Denver, CO.

Mann, L. (1979). On the trail of process. New York: Grune & Stratton, Inc.

Palincsar, A. S., & Brown, A. L. (1984). Reciprocal teaching of comprehension-fostering and monitoring activities. Cognition and Instruction, 1(2), 117-175.

Vlasova, T. A., & Pevzner, M. S. (Eds.). (1971). Children with temporary retardation in development. Moscow: Pedagogika.

Vygotsky, L. S. (1978). Mind in society: The development of higher psychological processes. (M. Cole, V. John-Steiner, S. Scribner, & E. Souberman, Eds. and trans.). Cambridge, MA: Harvard University Press.

Woodrow, H. (1917a). Practice and transference in normal and feeble-minded children. 1. Practice. Journal of Educational Psychology, 8, 85-96.

Woodrow, H. (1917b). Practice and transference in normal and feeble-minded children. 2. Transference. Journal of Educational Psychology, 8, 151-165.

Woodrow, H. (1946). The ability to learn. Psychological Review, 53, 147-158.

Wozniak, R. H. (1975). Psychology and education of the learning disabled child in the Soviet Union. In W. Cruikshank & D. P. Hallahan (Eds.), Research and theory in minimal cerebral dysfunction and learning disability. Syracuse, NY: Syracuse University Press.

Zabramna, S. D. (Ed.). (1971). _Otbor detei vo vspomogatel'nye shkoly_ (The selection of children for schools for the mentally retarded). Moscow: Prosveshchenie.

Table 1

Examples of Learning, Maintenance, and Transfer Items

Problem type	Pattern[a]	Sample Problem	Correct Answer
Original Learning	NN	N G O H P I Q J _ _ _ _	(R K S L)
	NINI	P Z U F Q Z V F _ _ _ _	(R Z W F)
Maintenance (Learned pattern types; new instantiations)			
Near Transfer (Learned relations and periodicities, but in new combinations)			
	NI	D V E V F V G V _ _ _ _	(H V I V)
	NNNN	V H D P W I E Q _ _ _ _	(X J F R)
Far Transfer (New relation, backward-next; or new periodicity, three letters)			
	BN	U C T D S E R F _ _ _ _	(Q G P H)
	NBNI	J P B X K O C X _ _ _ _	(L N D X)
	NIN	P A D Q A E R A _ _ _ _	(F S A G)
Very Far Transfer (Backward-next as well as next relations and "period" of two letters, but relations must be sought between strings of letters rather than within a string)			

Instructions:

Pretend that you are a spy. You want to send the message on top in a secret code that only your friends will understand. Someone has begun coding the message for you on the second line. Try to figure out the secret code and finish coding the message by filling in the blanks with the letters that follow the code.

S I X S H I P S G O N E

T H Y R I H Q R _ _ _ _ (H N O D)

[a] The letters themselves in the pattern notations refer to the alphabetic relations (i.e., N = next, I = identity, B = backward-next). The number of letters in each pattern notation equals the period.

Table 2

Multiple Regression Summary Table for Matrices Task

Dependent Variable	Independent Variable	Correlation r	Multiple R	Increment in R^2
Training	Information	-.439*	.439	.193*
	Coding	-.043	.587	.152*
Transfer	Information	-.389*	.389	.151*
Residual gain	Estimated IQ	.485*	.485	.235*
	Ravens	.472*	.608	.135*
	Training	-.605*	.770	.224*
	Transfer	-.598*	.876	.173*
	Far Transfer	-.698*	.884	.014

Table 3

Multiple Regression Summary Table for Series Completion Task

Dependent Variable	Independent Variable	Correlation r	Multiple R	Increment in R^2
Training	Block Design	-.476*	.476	.227*
	Vocabulary	-.427*	.677	.153*
Transfer	Block Design	-.581*	.581	.338*
	Animal House	-.479	.641	.073
Residual gain	Estimated IQ	.521*	.521	.272*
	Ravens	.352	.578	.062
	Training	-.461*	.595	.020
	Transfer	-.693*	.745	.221*
	Far Transfer	-.558*	.745	.000

Table 4

The Acquisition of Question-Asking by a Weak Seventh Grade Student (Charles)

Day 1:

TEXT: The water moccasin, somewhat longer than the copperhead, is found in the southeastern states. It lives in swampy regions. It belongs, as do also the copperhead and the rattlesnakes, to a group of poisonous snakes called pit vipers. They have pits between their eyes and their nostrils which, because they are sensitive to heat, help the snakes tell when they are near a warm-blooded animal. Another name for the water moccasin is "cottonmouth." This name comes from the white lining of the snake's mouth.

1) S: What is found in the southeastern snakes, also the copperhead, rattlesnakes, vipers - they have. I'm not doing this right.
2) T: All right. Do you want to know about the pit vipers?
3) S: Yeah.
4) T: What would be a good question about the pit vipers that starts with the word "why?"
5) S: (No response)
6) T: How about, "Why are the snakes called pit vipers?"
7) S: Why do they want to know that they are called pit vipers?
8) T: Try it again.
9) S: Why do they, pit vipers in a pit?
10) T: How about, "Why do they call the snakes pit vipers?"
11) S: Why do they call the snakes pit vipers?
12) T: There you go! Good for you.

Day 4:

TEXT: Spinner's mate is much smaller than she, and his body is dull brown. He spends most of his time setting at one side of her web.

13) S: (No question)
14) T: What's this paragraph about?
15) S: Spinner's mate. How do spinner's mate...
16) T: That's good. Keep going.
17) S: How do spinner's mate is smaller than...How am I going to say that?
18) T: Take your time with it. You want to ask a question about spinner's mate and what he does, beginning with the word "how."
19) S: How do they spend most of his time sitting?
20) T: You're very close. The question would be, "How does spinner's mate spend most of his time?" Now, you ask it.
21) S: How does spinner's mate spend most of his time?

Table 4 (cont'd.)

Day 7:

TEXT: Perhaps you are wondering where the lava and other volcanic products come from. Deep within our earth there are pockets of molten rock called magma. Forced upward in part by gas pressure, this molten rock continually tries to reach the surface. Eventually—by means of cracks in the crustal rocks or some similar zone of weakness—the magma may break out of the ground. It then flows from the vent as lava, or spews skyward as dense clouds of lava particles.

22) S: How does the pressure from below push the mass of hot rock against the opening? Is that it?
23) T: Not quite. Start your question with, "What happens when?"
24) S: What happens when the pressure from below pushes the mass of hot rock against the opening?
25) T: Good for you! Good job.

Day 11:

TEXT: One of the most interesting of the insect-eating plants is the Venus's flytrap. This plant lives in only one small area of the world -- the coastal marshes of North and South Carolina. The Venus's flytrap doesn't look unusual. Its habits, however, make it truly a plant wonder.

26) S: What is the most interesting of the insect eating plants, and where do the plants live at?
27) T: Two excellent questions! They are both clear and important questions. Ask us one at a time now.

Day 15:

TEXT: Scientists also come to the South Pole to study the strange lights that glow overhead during the Antarctic night. (It's a cold and lonely world for the few hardy people who "winter over" the polar night.) These "southern lights" are caused by the Earth acting like a magnet on electrical particles in the air. They are clues that may help us understand the Earth's core and the upper edges of its blanket of air.

28) S: Why do scientists come to the south pole to study?
29) T: Excellent question! That is what this paragraph is all about.

Table 5

Improvement in Question-Asking by a More Competent Seventh Grade Student (Sara)

Day 2:

TEXT: HOW CAN SNAKES BE SO FLEXIBLE?

> The snake's skeleton and parts of its body are very flexible--something like a rubber hose with bones. A snake's backbone can have as many as 300 vertebrae, almost ten times as many as a human's. These vertebrae are connected by loose and rubbery tissues that allow easy movement. Because of this bendable, twistable spinal construction, a snake can turn its body in almost any direction at almost any point.

1) S: Snakes' backbones can have as many as 300 vertebrates - almost ___ times as many as humans.
2) T: Not a bad beginning, but I would consider that a question about a detail. Try to avoid "fill in the blank" questions. See if next time y u can fine a main idea question and begin your question with a question word - how, why, when....

Day 3:

TEXT: There are snakes in nearly all parts of the world. Some snakes prefer warm, arid desert areas. Others prefer leafy forests, fields, and woodlands. Some stay in areas near water and are fine swimmers. Then there are several varieties that live all their lives in the sea.

3) S: Can snakes live their whole lives in seas?
4) T: See if you can ask a question using your own words.

Day 4:

TEXT: The other kind of camel--the one with two humps--is the Bactrian. Its home country is the Gobi Desert of northeastern Asia. The Bactrial has shorter legs and longer wool than the one-humped camel. It also has stronger, more rugged feet. This is important because instead of having sand to walk on, the .actrian camels live in rough and rocky parts of the world.

5) S: Where is the Bactrian found?
6) T: Good for you.

Table 5 (cont'd.)

Day 6:

TEXT: When most full-grown spiders want to travel, they have to walk on their eight legs. But some small kinds of spiders, and many young ones, use an easier way. They climb up on bushes, fence posts, or weed stems and spin streamers of silk. When the wind catches the silk and blows it away, each spider tightly holds onto his own streamer. The silk streamer carries him through the air as if it were a parachute or a balloon.

7) S: I think I have another. When it's traveling, what do they compare the spider to?
8) T: An interesting question.

Day 11:

TEXT: The young caterpillar's first meal is its own eggshell. Then it eats a leaf and each day eats more and more food. After a few days. the caterpillar becomes too large for its skin. A new skin forms beneath the first one, the old skin comes open and, like a snake, the caterpillar wriggles its way out of the split skin. Then the caterpillar goes on eating leaves or other kinds of food. When the new skin becomes too tight for the growing body. it again splits and comes off. By then the caterpillar is covered by another skin. This eating and shedding goes on for several weeks. The old skin may be replaced by a new one four or five times. Each time the skin is shed, the size and color of the caterpillar change.

9) S: Why does the caterpillar's skin split?
10) T: Excellent question. That was the point of the entire paragraph.

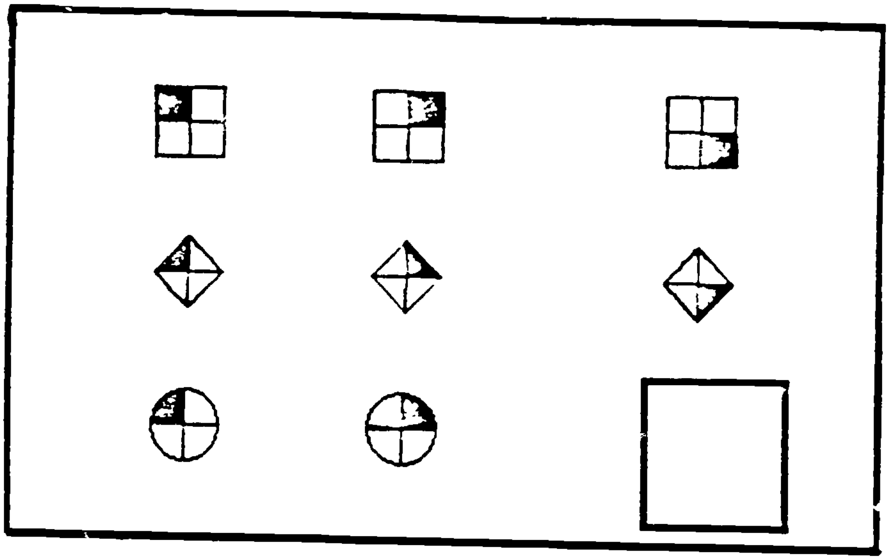
Figure Captions

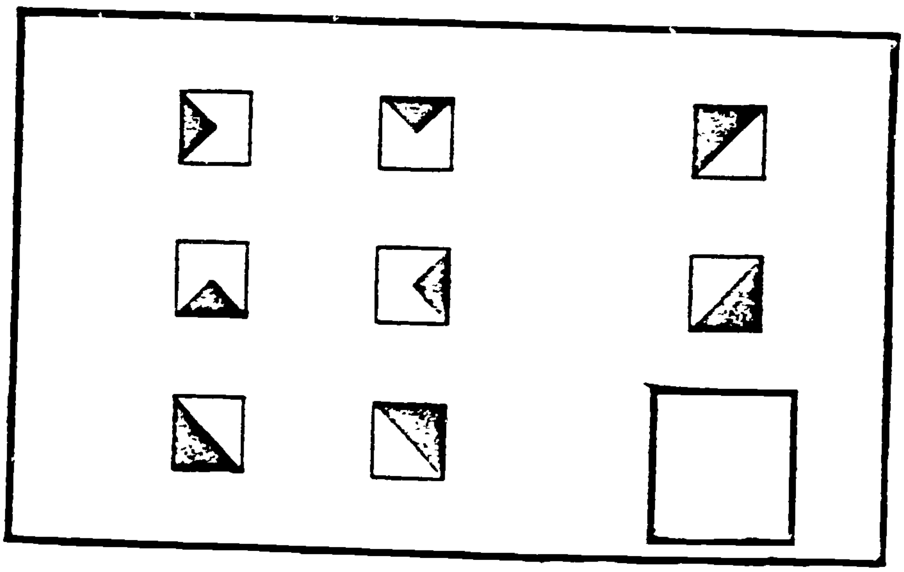
Figure 1. Examples of the learning problems used in the matrices study. The top panel contains a rotation problem, the middle panel an inposition problem, and the bottom panel a subtraction problem.

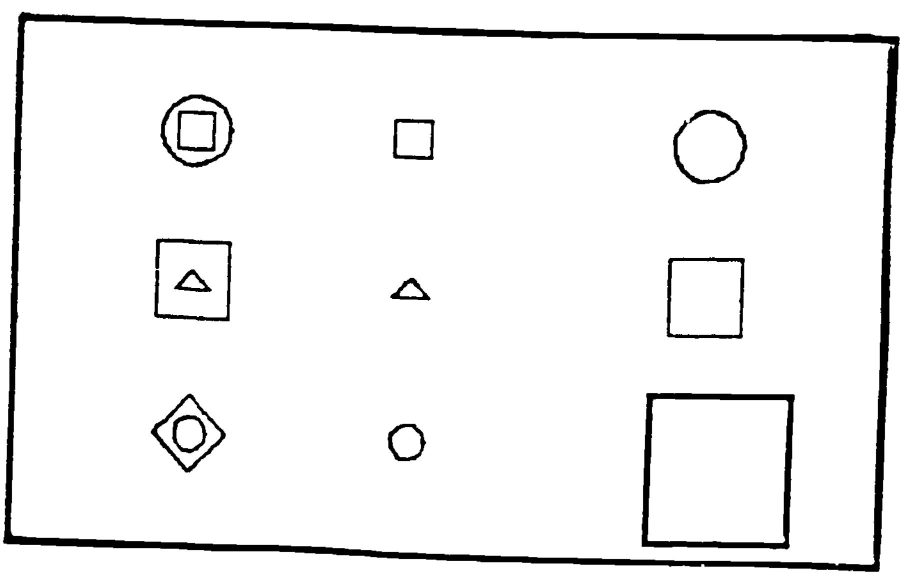
Figure 2. Examples of the transfer problems used in the matrices study. The top panel contians a rotation plus subtraction problem, the bottom panel a rotation plus imposition problem.

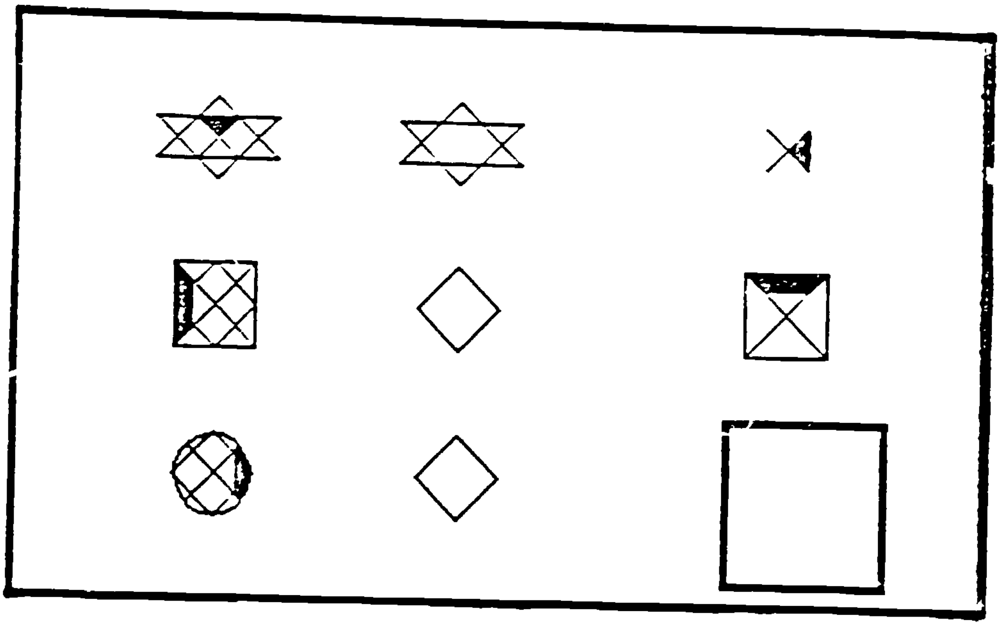
Figure 3. A sample hint sequence for rotation problem.

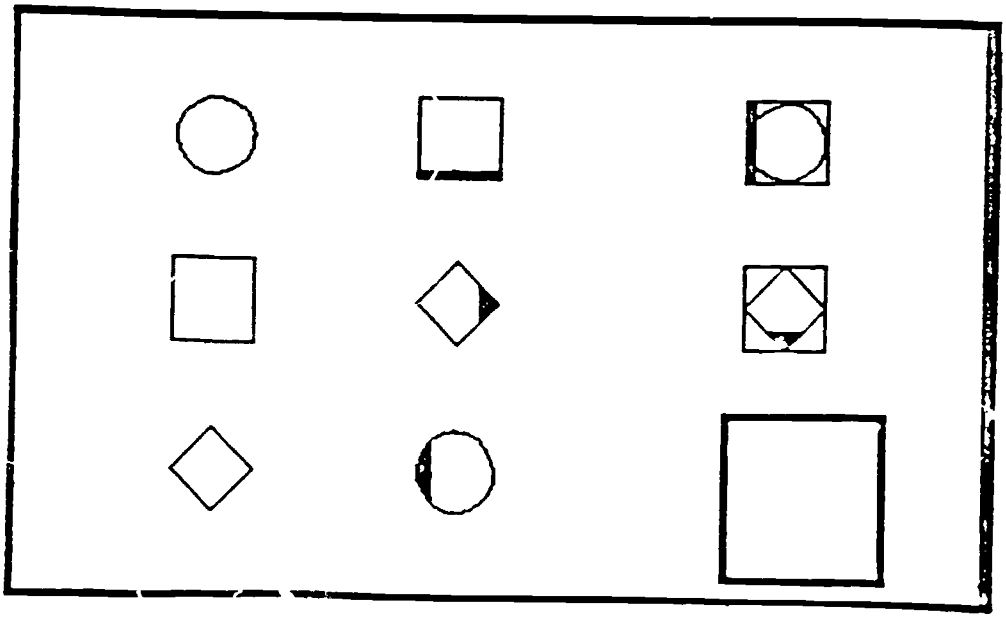
Figure 4. The mean number of prompts required on the transfer problems as a function of ability and transfer distance.











HINT 1: "THIS PROBLEM IS CALLED A TURNING PROBLEM. THINK ABOUT WHY IT MIGHT BE CALLED THAT...DO YOU KNOW HOW TO SOLVE THE PROBLEM NOW OR DO YOU WANT ANOTHER HINT?

HINT 2: "THIS IS ROW 1. PUT PICTURE 1 IN THE PRACTICE BOX. TOUCH IN. TOUCH THE PICTURE. NOW TRY TO MAKE THE PICTURE LOOK LIKE THE SECOND PICTURE. (IF SUCCESSFUL) YOU DID IT. NOW MAKE IT LOOK LIKE THE LAST PICTURE." (IF CHILD CANNOT MAKE PICTURE 3, PLATO WILL GIVE HINT 2A.)

HINT 2A: "THIS IS ROW 1. THIS IS PICTURE 1. WATCH HOW IT TURNS. WATCH AGAIN. NOW YOU DO IT." (IF CHILD CANNOT REPEAT THE ABOVE DEMONSTRATION, PLATO WILL GIVE HINT 2B.)

HINT 2B: "THIS IS ROW 1. LET'S TRY TO MAKE THE LAST PICTURE IN THE ROW. PUT PICTURE 1 IN THE PRACTICE BOX. TOUCH ^ . TOUCH ^ AGAIN. GOOD. YOU HAVE MADE THE LAST PICTURE IN ROW 1. NOW TRY TO MAKE THE MISSING PICTURE."

HINT 3: "NOW LET'S LOOK AT ROW 2. PUT PICTURE 1 OF ROW 2 IN THE PRACTICE BOX. NOW MAKE IT LOOK LIKE PICTURE 2. (IF CHILD DOES NOT RESPOND CORRECTLY PLATO WILL DISPLAY TOUCH ^ .) YOU DID IT. NOW MAKE THE PICTURE IN THE PRACTICE BOX LOOK LIKE THE LAST PICTURE IN ROW 2. NOW TRY THE PROBLEM AGAIN." (IF CHILD CANNOT MAKE PICTURE 3 PLATO WILL GIVE HINT 3A.)

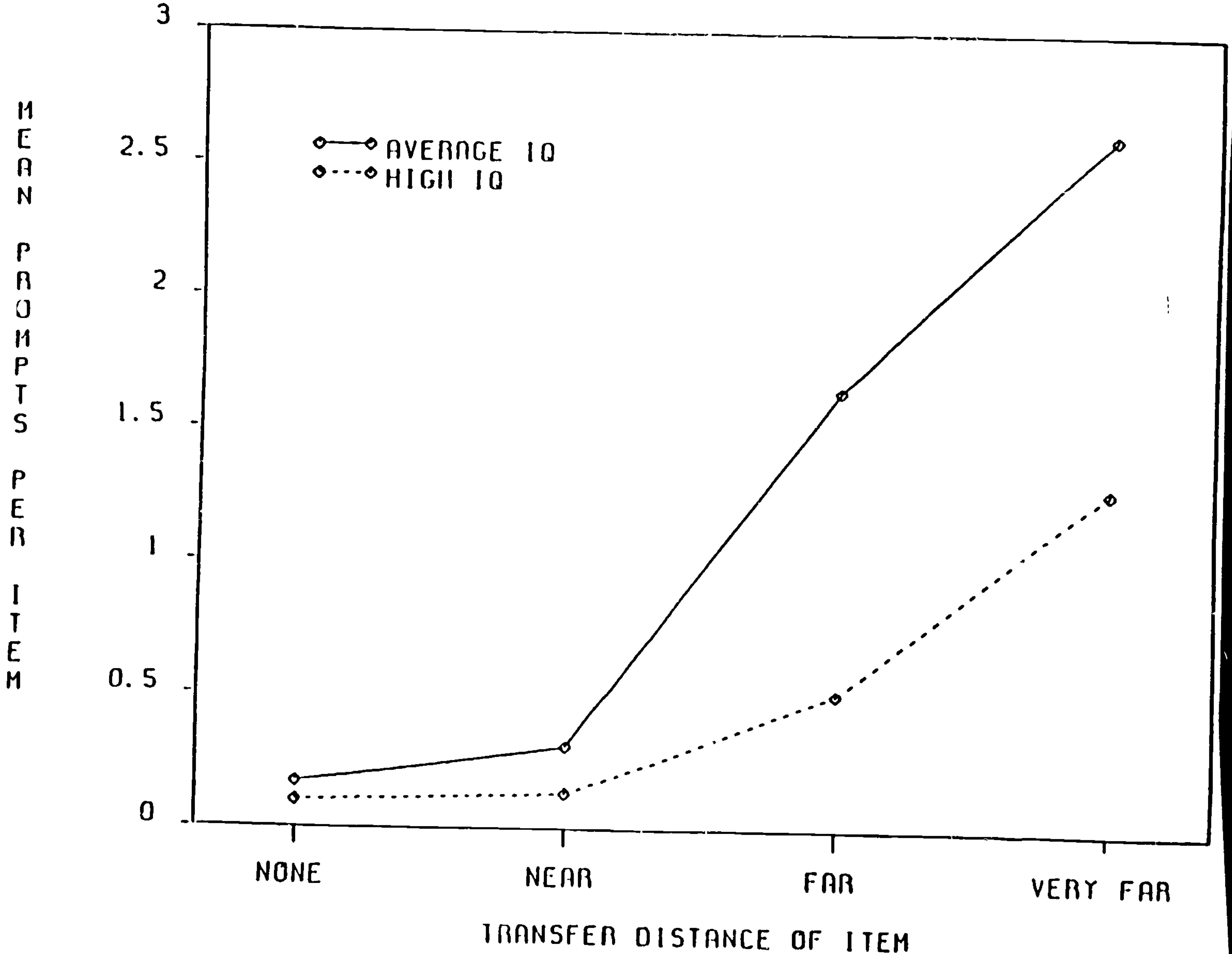
HINT 3A: "TOUCH ^ . TOUCH ^ AGAIN."

HINT 4: "YOU USED THE TURNING RULE TO MAKE THE LAST PICTURE IN ROWS 1 AND 2. THE LAST PICTURE IN ROW 3 IS MISSING. TRY TO USE THE SAME RULE TO MAKE THE MISSING PICTURE IN ROW 3." (IF CHILD CANNOT DO SO, PLATO WILL GIVE HINT 4A.)

HINT 4A: "THIS IS THE SHAPE YOU WORK WITH. (PLATO DISPLAYS APPROPRIATE SHAPE.) PUT IT IN THE PRACTICE BOX. TOUCH THE FIRST PICTURE IN ROW 3. NOW TOUCH ^ . NOW TOUCH ^ AGAIN. THAT IS CORRECT. TOUCH DONE."

EXPLANATION (GIVEN WITH EVERY ORIGINAL LEARNING PROBLEM): "GOOD. LOOK AT ALL THREE ROWS. THE TURNING RULE IS USED IN EACH ROW. AND YOU USED THE TURNING RULE TO MAKE THE MISSING PICTURE. YOU TURNED PICTURE 1 TO GET PICTURE 2. THEN YOU TURNED PICTURE 2 TO GET PICTURE 3.

IN ORIGINAL LEARNING, THE CHILD CONTINUES TO SOLVE ROTATION PROBLEMS UNTIL SHE CAN DO TWO PROBLEMS IN A ROW WITHOUT ANY HINTS. THEN PLATO WILL MOVE AHEAD TO THE FIRST IMPOSITION PROBLEM.


3
2.5
2
1.5
1
0.5
0
MEAN PROMPTS PER ITEM
AVERAGE IQ
HIGH IQ
NONE
NEAR
FAR
VERY FAR
TRANSFER DISTANCE OF ITEM