DOCUMENT RESUME

ED 079 373 TM 002 972

AUTHOR	Babad, Elisha Y.; Budoff, Milton
TITLE	Sensitivity and Validity of Learning Potential Measurement in Three Levels of Ability.
INSTITUTION	Research Inst. for Educational Problems, Cambridge, Mass.
SPONS AGENCY	Bureau of Education for the Handicapped (DHEW/OE), Washington, D.C. Div. of Research.; National Inst. of Mental Health, Rockville, Md.
PUB DATE	73
GRANT	OEG-0-8-080506-4597-(607)
NOTE	40p.; Paper presented at American Educational Research Association Meeting (New Orleans, Louisiana, February 25-March 1, 1973)
EDRS PRICE	MF-$0.65 HC-$3.29
DESCRIPTORS	*Achievement Gains; *Cognitive Ability; Comparative Analysis; *Intelligence Tests; Problem Solving; Technical Reports; *Testing; *Test Validity; Test Wiseness

ABSTRACT

In learning potential (LP) tests, intelligence is measured by repeated administrations of reasoning tasks, with interpolated training in problem-relevant strategies. In a comparison of the differential sensitivity and validity of LP and IQ measures, subjects were divided into three IQ groups: bright normal, dull-to-average, and subnormal (EMR). In the first study, both low IQ groups gained more from the coaching than did the bright group. In the second study, the concurrent validity of the LP test was equal to that of IQ in the entire sample and in the bright group, but superior to IQ in the low, disadvantaged groups. (Author)

SENSITIVITY AND VALIDITY OF LEARNING POTENTIAL MEASUREMENT IN THREE LEVELS OF ABILITY

Elisha Y. Babad and Milton Budoff

Research Institute for Educational Problems

Cambridge, Massachusetts

ABSTRACT

In the learning potential assessment procedures, intelligence is measured by repeated administrations of reasoning tasks, with training on problem-relevant strategies interpolated between administrations. These procedures are hypothesized to be more sensitive than traditional IQ tests in tapping the (yet unrealized) intellectual potential of disadvantaged children. The validity of the Series Learning Potential Test (SLPT) in bright normal, dull-to-average, and subnormal (educable mentally retarded) samples, was examined in two studies. In the first study, the SLPT was administered three times, the training in problem-relevant strategies interpolated following the second administration, to separate the effects of practice and coaching. Both low-IQ groups gained more than the high-IQ group from the training, and the dull group gained more than the other groups from repeated administrations without training. In the second study, the validity of the SLPT in predicting teacher ratings of school achievement was compared to the validity of a group IQ test. While the predictive powers of the SLPT and IQ scores were of the same magnitude for the entire sample and for the bright group, the SLPT was superior to IQ in the dull-to-average and the subnormal (EMR) groups. In both studies, substantial proportions of EMR subjects reached the average reasoning level of their non-retarded peers following the short training session

SENSITIVITY AND VALIDITY OF LEARNING POTENTIAL MEASUREMENT IN THREE LEVELS OF ABILITY[1]

Elisha Y. Babad[2] and Milton Budoff[3]

Research Institute for Educational Problems

Children from poor and/or nonwhite homes tend to score at below average levels on tests which purport to measure intelligence. Jensen (1969) has argued that the mean difference of 15 IQ points between white and black groups represents a real difference in inborn general ability. However, the IQ difference has also been frequently explained by the handicaps that poor and/or nonwhite children bring to the testing situation. They are fearful of the testing process, expect to do poorly, are often insensitive to speed requirements, and are unfamiliar with the problem contents.

IQ tests measure the degree to which children have spontaneously acquired, from their natural environment, the skills and knowledge which cumulatively predict academic school success. The plausible assumption is made that a child who learned informally prior to entering school will continue to learn -- formally and informally -- in and out of school. Children from non-middle-class homes who do not have an equal and frequent access to school-preparatory experiences, tend to score poorly on IQ tests, and are often viewed as "less intelligent."

Yet, many of these same low-IQ children are competent problem-solvers in their non-school environment, having mastered the skills, knowledge, and strategies necessary to maintain a successful adjustment. In other words, these children can learn and profit from relevant experiences more successfully than their IQ scores and school achievements indicate. In Hunt's (1961) terms, this difference in competence may represent the problem of the match between the environmental demands of the school and its tests, and the child's existing schemata in his familiar world. Assessment procedures must be developed which will optimize the match and result in more culture-fair measurement of general ability.

Davis and Eels (1951) developed problem games that were deemed relevant to the experiences of low-SES children, but found no change in the gap between the different social-class groups. This may have been due to the demand that the child go about solving the problems as middle-class psychologists thought he should. Changing the contents of the problems was not sufficient to enable the lower-class children to narrow the gap between their problem-solving styles and those expected by the middle-class school.

Budoff (1970) has described a learning potential (LP) procedure for assessing ability among low-IQ, low-SES children, based on a process-oriented conceptualization of intelligence. No new definition of intelligence or thinking is introduced in this conception. Instead, the focus is on the <u>acquisition</u> of reasoning and problem-solving strategies. The major concept in the LP paradigm is children's trainability, i.e. their ability to improve performance on reasoning problems following a systematic learning experience. Non-reasoning tasks, such as vocabulary and general information tests, are not used in LP tests, since they are based on specific prior experiences that are non-equatable for middle- and low-SES groups. Thus, even the potentially able culturally deprived child is left at a disadvantage when tested with a standard IQ test. In the LP paradigm reasoning is incorporated as the core of intelligence. The reasoning tasks are administered in a "test-coach-test" sequence, that compensates for the absence of spontaneous acquisition by low-SES children. The pretest reflects the subject's present level of functioning, and his existing ability to work with the problems. The training session provides the deprived child with those preparatory experiences necessary for solving the reasoning problems. In this session he familiarizes himself with the demands of the task, and receives sufficient opportunity to learn and successfully apply relevant strategies to the problems at hand. His post-test performance reflects the child's ability under optimal conditions. By this time all subjects are familiarized with the task and its demands, have had success in solving similar problems, and have had the opportunity for learning the usage and application of relevant strategies.

The children's learning in the training session is not mechanistic. That is, they learn a set of strategies, toegether with considerations necessary for adequate application of these strategies . It is assumed that the children's ability to learn how to think under such optimal conditions is indicative of a broader ability to learn within other contexts, if appropriately taught. It is also assumed that middle-class children do not stand to gain much from the training session, as they have had enough prior problem-solving experiences to form adequate approaches to the task on the pretest. However, the test should differentiate among the low-SES children more sharply than the traditional IQ tests, identifying those who are potentially able but who have always failed because of deficiencies in their natural environment.

The problems employed in the LP tests fall within Jensen's (1969) Level II category (conceptual learning). According to Jensen, there is but little overlap between the Level II curves of lower - and higher - SES groups. Jensen's theory is based upon "one-shot" intelligence testing. In the authors' opinion, such testing procedures underestimate the potential Level II ability of substantial proportions of disadvantaged children. Therefore, one would expect to find greater overlap between the curves of lower - and higher - SES groups on LP post-tests, than on LP pretests, and than on other one-shot tests. We do not assume complete overlap between the Level II curves. Using LP tests, we merely attempt to identify the numerous false positives among the disadvantaged.

Three non-verbal tasks have been employed in assessment of LP: an altered version of the Kohs Block Designs (Budoff, 1969), Raven's Progressive Matrices (Budoff, 1970), and the recently constructed Series Learning Potential Test (SLPT, Babad and Budoff, 1971). The SLPT, used in the present study, is a group test for the primary and elementary grades, its major task is the completion of series. The two equivalent forms of the test contains 65 multiple-choice items, each of which 50 are series problems, and the remaining are double-classification matrices. Forty series (and five matrices) are presented in simple pictures and the remaining items are presented in geometric figures. The coaching booklet contains 17 picture-series problems, all of which are solved by the entire class during the coaching session. Besides evaluating both improvement and final level of ability in solving picture-series problems, the test is designed to evaluate generalization from pictorial to geometric stimuli. Geometric stimuli are not used in coaching, but the strategies applied to them are the same as those taught in the coaching session. The test is also designed to examine transfer from series to matrix problems, where the problems are conceptually similar, but the strategies must be modified to adequately apply to the matrices. Five change dimensions are used in the entire test: object, size, color, position and filling of geometric objects. A given series contains no more than four dimensions changing concurrently. Unlike Raven's matrices, where some of the harder items involve abstract algebraic concepts, all SLPT concepts are simple and rhythmic. The level of difficulty of a given item is a function of the number of concepts it contains and the asymmetry between them. The different series within a given item are always independent of each other.

A sample multiple-choice item from the SLPT is presented in Figure 1.

Insert Figure 1

This item consists of three independent series (object, color, size), changing asymmetrically. The schematic presentation of this item is as follows:

1)	Object series:	TB	AP	FL	TB	AP	FL	___	AP	FL	TB
2)	Color series:	W	W	B	B	W	W	___	B	W	W
3)	Size series:	L	S	L	S	L	S	___	S	L	S

To solve this problem most efficiently, the subject must separate the three series, solve each independently of the others, and then converge his three solutions onto one correct answer. Since the change on the different dimensions is asymmetrical and the dimensions independent of each other, it is inefficient to try to solve the problem by checking all the variables at once.

Five strategies are taught and practiceu in the coaching session:

1) Using rhythm as a mental organizational or chunking aid in order to form the concept of each series. Concepts are thus arrived at by "singing the tune" of each series, without necessitating further verbalization.

2) Separating the series and looking for an independent solution to each change dimension.

3) Reducing the memory load by eliminating the wrong choices for each solution.

4) Reversing the direction in which the tune is sung, when the location of the blank space calls for such reversal.

5) Identifying the starting point of the series, the first picture of which is in mid-tune.

The strategies for handling the stimuli are essentially problem solving rules combining operations of deductive and inductive thinking. Each item is first analyzed into its components, independent concepts are separately formed, and only then are they converged into a synthesized solution.

Rhythm is used here as a mental organizational aid, combined with a strategy for reducing memory load by elimination of wrong choices. Pribram (1967) has already noted that appropriate chunking improves the ability to organize stimuli and solve problems, even to the extent for compensating for frontal lobe lesions. The reasoning process in this context is thus translatable as the formation of efficient ways of approaching the problems and of organizing the stimuli.

Solution of SLPT problems involves concrete-operational non-formal thougnt. Since Piaget, it is generally accepted that thinking in the primary and elementary levels consists primarily of forming and using appropriate ways of directly handling stimuli, and that these approaches are acquired through experiential interaction with the stimulus environment. It is only at a later stage that the abstractness and complexity of the concepts themselves come to the fore as the core of the thinking act.

To appropriately apply the strategies to SLPT problems, the child is required to have previous capabilities in perceptual discrimination, directionality, labeling, reversing, and ordering. these are not taught during the LP coaching session, but are presumed to exist in the child's repertoire, to be utilized as the basis for application of relevant rules and principles. These rules and principles are algorithmic, but in teaching them we do not utilize an S-R approach to learning. We teach how to approach the problems, not to perform automatic learned responses. Unlike such algorithms as extracting square roots, the SLPT strategies cannot be used without a real understanding of the task and its demands. Thus, the LP scores reflect the degree to which the children are operational within the given framework. Bereiter has advocated an educational approach (for dealing with problems of sequence and organization) utilizing the teaching of rules and principles to disadvantaged preschoolers, (Bereiter, 1970, Bereiter and Englemann, 1966.).

Although the SLPT contains a training session, it is basically a test of intelligence, not a training program. The didactic approach used in the SLPT differs in the specificity of its teaching from that advocated and used in a Piagetian curriculum. Kamii (1970) states that "the process of hypothetico-deductive thinking, not rules and their applications, is at the center of Piagetian interest." (p. 30) She translates this to mean that "the imposition of rules is a less efficient way to teach than influencing the development of underlying cognitive processes that will eventually enable the child to construct his own rules." The wide-ranged and time-consuming basic training advocated by Kamii is impracticable in a measurement paradigm. Since an intelligence test can allow only a limited amount of training, we tried to develop a set of strategies that is parsimonious and efficient, but still making it possible to evaluate the extent to which the child can operate with given rules and principles. Unlike Kamii, we do not ask the child to verbally explain his manipulations of the materials, but test his operational ability in a situation where the correct answer cannot be reached without true understanding of the underlying rules and principles.

Ortar (1960) reported that coaching on a non-verbal reasoning task improved the test's validity in predicting school-type achievements. Based on this and other findings she concluded that intelligence-tests constructed in a test-coach-test paradigm may be more valid than traditional tests, particularly for use wi s‿.vantaged children.

More recently, Guinagh (1971), Jacobs and Vandeventer (1971), Linn (1971) and Feuerstein in Israel (1968) have experimented with coaching on double-classification problems. However, they did not investigate the sensitivity and validity of their measures in different social-class and ability groups, and did not compare the predictive powers of these tests to those of traditional measures of intelligence in actual educational settings.

Budoff (1968, 1969, 1970) found that IQ-defined dull-to-average and educable mentally retarded (EMR) subjects display three patterns of response on LP tests. Some subjects, "high scorers", perform at the level of their higher IQ peers on the pretest. Others, "gainers", perform at a poor level prior to training, but improve their performance on the post-test, often reaching the level of their higher IQ agemates. Only their pretest scores are consonant with their IQ scores. The third group, "non-gainers", perform poorly on both pretest and post-test, showing no profit from tuition.

While these three groups typically cannot be distinguished from each other on the basis of WISC or SB IQs, the high scorers and gainers systematically differ from the low-LP group on a variety of dimensions, that range from non-verbal problem-solving t sks to motivational, attitudinal, and personality variables (Budoff, ·:8 1969, 1970). Budoff, Meskin, and Harrison (1971) tested the acquisition of principles of electricity in a manipulative science program, in groups of special and regular class students. They found that LP status best differentiated levels of attainment following this course, while neither IQ nor special versus regular class placement could distinguish levels of achievement in the course.

The general hypothesis of the two studies reported in this paper is that the SLPT will show a high level of differentiation among low-IQ, low-SES subjects. We expect the SLPT to have higher predictive validity and greater sensitivity in identifying potential ability for these groups than does a standard group test of intelligence. Such difference in sensitivity and validity is not expected for high IQ-groups. In both studies subjects were divided into three IQ groups: 1) bright, normal, middle-class children (IQs above 100); 2) blue-collar, dull-to-average children (IQ 80 to 99); and 3) poor, subnormal, educable mentally retarded (EMR) children (IQs below 80). Five specific hypotheses will be tested:

1) Low-IQ children will gain from repeated, uncoached exposures to the SLPT more than will high-IQ children (practice effect, to be tested in Study I).

2) Low-IQ children will gain from training on problem-relevant strategies more than will high-IQ children (coaching effect, to be tested in Study I).

3) SLPT scores will predict school achievements of low-IQ groups better than group IQ scores will. This will not be true for the high-IQ group. (To be tested in Study II).

4) The power of the SLPT to predict school achievements will increase from pretest to post-test for the low-IQ groups, but not for the high-IQ group. (To be tested in Study II).

5) Following the training, substantial proportions of IQ-defined EMR subjects will reach the mean pretest score of their non-retarded peers. (To be tested in both studies). If these subjects can reach the reasoning level of non-retarded children following a very short tuition, they must be _educationally_, not _mentally_ retarded.

STUDY I

The SLPT was administered three times to groups of bright, dull-to-average, and subnormal (EMR) children in the middle elementary school years. The coaching session was interpolated following the second test administration. Three scores were calculated for each subject: practice score ($SLPT_2$ minus $SLPT_1$), coaching score ($SLPT_3$ minus $SLPT_2$), and a combined gain score ($SLPT_3$ minus $SLPT_1$).

Method

Subjects

Subjects were 126 (58 boys, 68 girls) white children in third, fourth, and fifth grades of several New England schools. The subjects were divided (into three groups) according to their IQ scores on the Test of General Ability (Flanagan, 1960):

1) bright normal ($\underline{N}$=64, 21 boys and 43 girls, mean IQ of 113 $\pm$12, IQ$\geq$100, and predominantly from middle-class, suburban homes);
2) dull-to-average ($\underline{N}$=37, 17 boys and 20 girls, mean IQ of 85 $\pm$7, 80$\leq$IQ$\leq$99, predominantly from blue-collar homes in an inner city district);
3) subnormal (EMR), drawn from special classes for the mentally retarded ($\underline{N}$=25, 20 boys and 5 girls, mean IQ of 68 $\pm$7, IQ$<$80, from blue-collar homes in an inner city district).

There was no indication in the school records of organic brain pathology in any subject.

Instruments

In this study, the Test of General Ability (TOGA, Flanagan, 1960) was used to divide the subjects into three IQ groups, while the three SLPT (Babad and Budoff, 1971) scores served as the dependent variables.

The TOGA is a group-IQ test, consisting of verbal and reasoning parts. Both parts are multiple-choice tests, with pictorial stimuli in the verbal part, and abstract symbols in the reasoning part. Three scores are derived from the TOGA: Verbal IQ, Reasoning IQ, and Total IQ.

The SLPT, as described above, consists of two equivalent 65-item multiple choice tests. Form A is used for the pretest, and form B for the post-test. Each form is administered as a power test and usually takes about 30-35 minutes to complete. The coaching booklet consists of 17 items, all of which are solved together by the entire class. One student is invited to solve each problem aloud, and the rest of the class silently follows in their booklets. The tester explains the various strategies as they become relevant to specific problems. The children are then encouraged to apply these strategies to further items. The tester tries to increase class participation, giving the less active and less able children a chance for public success by having them solve some of the problems aloud in front of the class. All 17 problems are successfully solved by the end of the coaching session. The coaching lasts for thirty to fifty minutes, depending on such factors, as class size, class level and the amount of participation.

Procedure

The study was conducted during four sessions in the spring of the school year. The SLPT was first administered in three sessions, with two-day intervals between sessions. All subjects received Form A (pretest) in the first session. In the second session, half of the subjects received Form A again, and the other half, Form B. For both groups the training session immediately followed. In the third session (all subjects received Form B (post-test)). The TOGA was administered several days after these three sessions. All tests were administered by experienced assistants, following standard instructions. In each class, all tests were administered by the same person.

Results

Practice gain scores, coaching gain scores, and combined gain scores were calculated for the three groups. The patterns of these scores are presented in Figure 2. The means and standard deviations of the three groups in initial and final testing are

Insert Figure 2

presented in Table 1.

Insert Table 1

To test the first and second hypotheses, separate one-way analyses of variance were computed for practice effect, coaching effects, and combined gains. These analyses were followed by *t*-tests between the three possible pairs of groups. The results of the analysis of variance and the subsequent t-tests are presented in Table 2.

Insert Table 2

As can be seen in Figure 2 and Tables 1 and 2, both the first and second hypotheses were confirmed. Repeated administrations without training resulted in differential score increments for the three groups. The dull-to-average group gained from practice significantly more than did the subnormal and the bright normal groups, while the difference in practice gains between the latter groups was non-significant. As to coaching effects, again we found differential increments in score, but the pattern of the groups differed from that found for practice gains. Both the subnormal and the dull-to-average groups gained more from training than did the bright normal group. The difference in coaching effect between the two low-IQ groups was not significant, but it should be noted that the highest gain was found for the subnormal group.

The analysis of variance of the combined gain scores again yielded a significant F-ratio indicating differential increments in score. Again, both low-IQ groups gained more than did the bright normal group, while the difference between the low-IQ groups was not significant.

Since half of the subjects received Form A of the SLPT in the second session, while the other half received Form B, the analyses of variance were recomputed for each half. All effects and patterns were similar, and the means of the groups were almost identical for each half.

Repeated administrations of the SLPT -- with and without training -- thus resulted in differential effects on the three groups. The disadvantaged, low-IQ children learned and profited from training on problem-relevant strategies more than did their middle-class, high-IQ peers, and the dull-to-average group gained as much from mere practice as well.

The relative spread of TOGA and SLPT scores of the three groups is also indicative of the sensitivity of the SLPT in the low-SES, low-IQ groups. While the standard deviation of IQ scores for the bright normal group (12) is almost twice as large as the standard deviations for both low-IQ groups (both 7 points), the picture is reversed for SLPT scores. Here the standard deviations of both low-IQ groups are twice as large as those of the bright normal group, both in initial and final testing. (See Table 1). The differential effect of training is further indicated in the changes of the SLPT standard deviations from initial to final testing. A slight shrinkage in standard deviation was found for the bright group, as compared with the 20% increase in the standard deviation of the subnormal group.

The possibility that the observed pattern of results was caused by a ceiling effect for the bright normal group in the post-test was explored by checking this group's performance in the post-test. The mean score (55.4) was more than two standard deviations below the ceiling (65) of the SLPT. The distribution of scores approximated the normal curve, with no indication of that skewedness which typically characterizes ceiling effects. This led us to rule out a ceiling effect as the major determinant of the reported patterns.

STUDY II

The validity of SLPT scores and TOGA IQs in predicting teacher-rated school success was compared for bright normal, dull-to-average, and subnormal (EMR) samples. The SLPT was predicted to be superior in validity to the TOGA in the low-IQ, but not in the high-IQ, range. The validity of the SLPT was expected to improve from pretest to post-test in the low-IQ, but not the high-IQ range.

Method

Subjects

The subjects of the first study were selected from separate schools, the EMR sample consisting of students in segregated special classes for the retarded. To equate samples, schools, classes, and teacher grades, a new sample was selected with a wide representative range of ability in each class. In this sample, all EMR subjects were integrated into regular classes.

Subjects were 207 white children in twelve third-grade classes of three elementary schools. All schools were located in a small New England town, with a predominantly white working-class population. The subjects were divided into three groups according to their TOGA IQ scores (Flanagan, 1960): bright normal (N=76, mean IQ of 113 ± 9, IQ$\leq$100); dull-to-average (N=95, mean IQ of 88 ± 6, 80$\leq$IQ$\leq$99); and subnormal (N=36, mean IQ of 72 ± 6.5 , IQ$<$80, classified in Massachusetts as EMR). There was no indication in the school records of organic brain pathology in any subject. Subjects for all three groups were found in each of the twelve classes.

Instruments

As in the first study, the SLPT (Babad and Budoff, 1971), and the TOGA (Flanagan, 1960) were used. School achievement was measured by teacher ratings, for reading, spelling, arithmetic, "general achievement," and "academic potential." The ratings were done on an eleven-point scale, corresponding to the letter grade system which the teachers are accustomed to use in that school system.

Procedure

The study was conducted in three sessions during the spring of the school year. In the first session the SLPT pretest (Form A) was administered to all students, followed by the standardized training. The teachers completed the rating scales during this period. The SLPT post-test was administered in the second session, three days after the first. The TOGA was administered several days later by the same tester.

Results

The relationships of LP and IQ variables to teacher ratings of school achievement are presented in Tables 3, 4, 5, and 6 for

Insert Tables 3,4,5, and 6

the entire sample, the subnormal, dull-to-average, and bright normal groups, respectively. The correlations of LP with achievement and IQ with achievement were of the same magnitude in the total sample (Table 3). As hypothesized, the LP predictions of achievement were higher than the IQ predictions for the subnormal and dull-to-average samples. (Tables 4 and 5). For the bright normal group there was almost no difference in predictive power. (Table 6). In fact, the post-test LP predictions for the two low-IQ groups were more than twice as large as those of IQ scores for these groups. Because of the restricted range, all subsample correlations decreased; but the predictability of IQ scores suffered far more from this restriction than did LP scores for the low-IQ groups. The third hypothesis was thus confirmed. The validity of SLPT was not inferior to that of TOGA IQ in the overall sample, and superior to IQ in the restricted low-IQ range.

As to the fourth hypothesis, the expected changes in validity of the SLPT from pretest to post-test were found (increase for the low-IQ groups, no increase for the high-IQ group), but the changes were small and inconclusive. The hypothesis was thus not confirmed by the data. One should also note the relative superiority of reasoning IQ predictions over verbal IQ predictions for both low-IQ groups, and the absence of such difference in the bright normal group.

One may view yet another strength of the training-based LP measurement, by determining the proportion of IQ-defined EMR subjects who attain, the mean pretest score of their non-retarded agemates. (following tuition). According to our trainability hypothesis, EMR subjects who reach the pre-training reasoning level of non-retardates (following a short training session), may have been falsely identified as mentally retarded. Proportions of EMR subjects falling above the means pretest score of non-retarded groups (dull-to-average and bright normal) were calculated for the samples in Study I and Study II, and are presented in Table 7.

Insert Table 7

The changes from pre- to post-training proportions are quite marked in the four comparisons in Table 7, indicating that following a short problem-relevant learning experience, substantial proportions of IQ-defined EMR subjects do attain the average (non-coached) reasoning level of their non-retarded peers. The fifth hypothesis was thus confirmed.

As seen in Table 7, the trainability effect was particularly striking in the second study, where the EMR sample was drawn from regular classes, rather than from segregated special classes for the mentally retarded. In that study, mean EMR post-test performance exceeded the mean pretest level of the dull-to-average group, despite the 16-point IQ difference between the groups. In their post-test performance, more than a third of the EMR subjects surpassed the mean pretest score of the bright group, despite a 41-point (!) difference between the average IQs of the two groups.

In Study II, we further compared the proportions of EMR subjects whose post-test scores were equal to, or above, the mean <u>post-test</u> score for non-retarded subjects, with the proportions of EMR subjects whose pretest scores equalled, or exceeded, the mean <u>pretest</u> scores of their agemates. Nineteen percent of the EMR subjects attained pretest scores at, or above, the mean pretest score of the dull-to-average group, and 28% of the EMR subjects' post-test scores were at, or above, the mean post-test level of the dull group. Compared to the bright normal group of Study II, the EMR proportions were 3% and 8% for the pre-and post-training scores. Thus, despite the absence of any IQ overlap between the groups, the relative positions of some EMR subjects improved even when compared to their coached peers.

Discussion

The notion that LP testing would show a higher level of differentiation among low-IQ, low SES subjects, than a standard test of intelligence would was borne out by the results. We found that utilization of the SLPT led to the discovery of a high level of trainability among low-IQ children, previously undetected by standard measures of intelligence. Disadvantaged dull-to-average and EMR groups profited from coaching on problem-relevant strategies more than did the bright normal group, while the dull-to-average group gained as much from practice as from coaching on the task. Substantial proportions of so-called EMR children reached the reasoning level of their brighter peers following a short training session on problem-relevant strategies. The observed sensitivity of LP measurement in the low-IQ range was further borne out by the superiority of the SLPT to TOGA IQ in predicting school achievements. While the predictive powers of both tests were of the same magnitude in the total sample and the bright normal group, the SLPT was a far better predictor than IQ in the dull-to-average and EMR groups.

To permit reasonable predictions concerning educability in a school setting, reasoning abilities of the child must be tested. These abilities are inevitably tainted by prior cultural experiences of the subjects. By providing all children with directed training on problem-relevant strategies, the LP paradigm seeks to bring all children to a more equalized starting point prior to testing, thereby giving high-SES children advantage over low-SES children in the post-training measurement. Indeed, we found a considerably greater overlap of scores among groups differing in IQ following training than there had been prior to training. Given the opportunity to become familiar with the demands of the task and to learn how to approach the problems, the lower-IQ children from blue-collar and poor backgrounds displayed considerably more competence, confidence, and sense of challenge on the task than they had previously. In Hunt's (1961) terms it could be said that, for the post-training measurement, there was a better match between the children's existing schemata and the demands of the situation. For the middle-class children, the "problem of the match" did not exist, and they performed as well on the pretest as on the post-test.

The increased overlap between reasoning curves of high- and low-SES subjects following coaching indicates that the Level II ability of at least portions of disadvantaged populations is greater than that posited by Jensen on the basis of on-shot tests. Children, who on the basis of their IQ and achievement scores, were considered to have a low capacity for Level II thinking, have now shown a greater capacity than predicted for successful performance at this level. Since it is now possible to identify from among disadvantaged, low-IQ children, those who are reachable and who have potential for functioning within a framework of Level II thinking, it would seem promising for educators to work differentially with these children, instead of applying undifferentiated programs to random samples of disadvantaged children.

In the present context, we have no conclusive statements to make about how best to teach high-potential, disadvantaged children. The SLPT training, utilizing the teaching of strategies, rules, and principles, was found successful in bringing disadvantaged children to a higher operational level of intelligence. Budoff, Meskin, and Harrison (1971), on the other hand, found equal success with learning via physical manipulation, requiring a minimum of verbalization and allowing the children to arrive at their own operational definitions of the relevant laws.

Judging by the success of the LP approach, all successful teaching programs would seem to profit by incorporating the following: the initial utilization of channels in which children are less deficient, the provision of the necessary preparatory learning experiences, and the creation of subjective feelings of mastery and success.

REFERENCES

Babad, E.Y. and Budoff, M. A preliminary manual of the Series Learning Potential Test. Studies in Learning Potential, 1971, Vol. 2.

Bereiter, C. and Engelman, S. Teaching the Disadvantaged Child in the Preschool, Engelwood Cliffs, N.J., Prentice Hall, 1966.

Bereiter, C. Educational Implications of Kohlberg's Cognitive-Developmental View, Interchange, 1970, 1, (1)

Budoff, M. Learning potential as a supplementary assessment procedure. Learning Disorders, 1968, 3, 295-343.

Budoff, M. Learning potential: A supplementary procedure for assessing the ability to reason. Seminars in Psychiatry, 1969, 1, 278-290.

Budoff, M. Learning potential hypothesis: Its effects upon teacher expectancies for disadvantaged students. Paper presented at the Conference on Teacher Expectancies, Georgia State University, Atlanta, Georgia, October, 1970.

Budoff, M., Meskin, J., and Harrison, R.H. Educational test of the learning potential hypothesis, American Journal of Mental Deficiency, 1971, 76, 159-169.

Eels, K., Davis, A., and Havighurst, R.J. et al. Intelligence and Cultural Differnces. Chicago, University of Chicago Press, 1951.

Feuerstein, R. The Learning potential assessment device, Mimeo. Jerusalem, Youth Aliyah Department of the Jewish Agency, 1968.

Flanagan. J.C. Test of General Ability: technical report. Chicago: Science Research Associates, 1960.

Guinagh, B.J. An experimental study of basic learning ability and intelligence in low-socio-economic-status children. Child Development, 1971, 42, 27-36.

Hunt, J. McV. Intelligence and Experience. New York, The Ronald Press Company, 1961.

Jacobs, P.J. and Vandeventer, M. The learning and transfer of double-classification skills by first graders. Child Development, 1971, 42, 149-159.

Jensen, A.R. How much can we boost IQ and scholastic achievement? Harvard Educational Review, 1969, 39, 1-123.

Kamii, C. An application of Piaget's theory to the conceptualization of a preschool curriculum. Mimeo, Ypsilanti Public Schools, Michigan, 1970.

Linn, M. The role of intelligence in children's response to instructions. Unpublished paper, School of Education, Stanford University, 1971.

Ortar, Improving test validity by coaching, Educational Research, 1960, 2, 137-142.

Pribram, K.H., and Tubbs, W.E. Short term memory, parsing, and the primate frontal cortex, Science, 1967, 1765-1767.

FOOTNOTES

[1] This research was supported by Grant OEG-0-8-080506-4597(607) from the Research Branch of the Bureau of the Handicapped, U.S. Office of Education, and Grant MH18553 from the National Institute of Mental Health. The authors wish to extend their warm gratitude to Kathleen T. Wessling, whose assistance helped carry out this study, and to Bina Presser, who helped put everything together.

[2] Dr. Babad is now at the School of Education, Hebrew University of Jerusalem, Israel.

[3] Requests for reprints should be sent to Dr. Milton Budoff, Director, Research Institute for Educational Problems, 12 Maple Avenue, Cambridge, Massachusetts, 02139.

Table 1. Means and Standard Deviations of SLPT and TOGA Scores of the Three Groups in Study 1

		$SLPT_1$ Initial Performance		$SLPT_3$ Final Performance		TOGA IQ	
Group	N	$\bar{X}$	SD	$\bar{X}$	SD	$\bar{X}$	SD
Bright Normal	64	51.9	5.0	55.4	4.7	113	12.0
Dull-to-Average	37	37.6	10.1	47.3	9.9	85	7.0
Subnormal (EMR)	25	26.4	10.7	35.0	12.3	68	7.0

Table 2. One way analyses of variance and subsequent *t*-tests of practice effects, coaching effects, and combined gain effects of the three groups in Study I.

	One-way ANOVA *df*=2,125	*t*-tests: bright vs. dull *df*=99	bright vs. sub-normal *df*=87	dull vs. sub-normal *df*=60
Practice effect	*F*=16.3 $P<.001$	*t*=3.33 $P<.001$	*t*=0.56 n.s	*t*=2.54 $P<.01$
Coaching effect	*F*=2.4 $P<.10$	*t*=2.26 $P<.025$	*t*=2.69 $P<.005$	*t*=0.54 n.s.
Combined gain effect	*F*=9.32 $P<.001$	*t*=4.03 $P<.001$	*t*=2.89 $P<.005$	*t*=1.03 n.s.

Table 3. Co-efficients of correlation of SLPT and TOGA scores with teacher ratings in the total sample ($\underline{N}$=207), Study II.

teacher ratings \ predictor	SLPT pretest	SLPT posttest	TOGA verbal	TOGA reasoning	TOGA IQ
Reading	.50	.50	.45	.49	.54
Spelling	.45	.49	.45	.47	.53
Arithmetic	.49	.53	.46	.50	.52
General Achievement	.51	.54	.49	.52	.56
Academic Potential	.49	.50	.43	.47	.50
Average Correlation	.49	.51	.46	.49	.53

Table 4. Co-efficients of correlation of SLPT and TOGA scores with teacher ratings in the subnormal (EMR) sample ($\underline{N}$=36), Study II.

teacher ratings \ predictor	SLPT pretest	SLPT postest	TOGA verbal	TOGA reasoning	TOGA IQ
Reading	.24	.28	-.01	.20	.12
Spelling	.26	.38	.19	.20	.20
Arithmetic	.35	.49	.18	.33	.18
General Achievement	.33	.40	.07	.25	.15
Academic Potential	.27	.22	.11	.01	.13
Average Correlation	.29	.35	.11	.20	.16

Table 5. Co-efficients of correlation of SLPT and TOGA scores with teacher ratings in the dull-to-average sample (N=95), Study II.

teacher ratings \ predictor	SLPT pretest	SLPT posttest	TOGA verbal	TOGA reasoning	TOGA IQ
Reading	.35	.35	.05	.15	.09
Spelling	.23	.26	.08	.17	.17
Arithmetic	.27	.29	.14	.16	.18
General Achievement	.31	.34	.08	.22	.19
Academic Potential	.35	.44	.04	.23	.11
Average Correlation	.30	.34	.08	.19	.15

Table 6. Co-efficients of correlation of SLPT and TOGA scores with teacher ratings in the bright normal sample (N=76), Study II.

teacher ratings \ predictor	SLPT pretest	SLPT posttest	TOGA verbal	TOGA reasoning	TOGA IQ
Reading	.35	.33	.27	.23	.29
Spelling	.32	.32	.26	.22	.33
Arithmetic	.43	.40	.26	.34	.32
General Achievement	.40	.38	.33	.26	.32
Academic Potential	.37	.35	.25	.25	.29
Average Correlation	.37	.35	.27	.25	.31

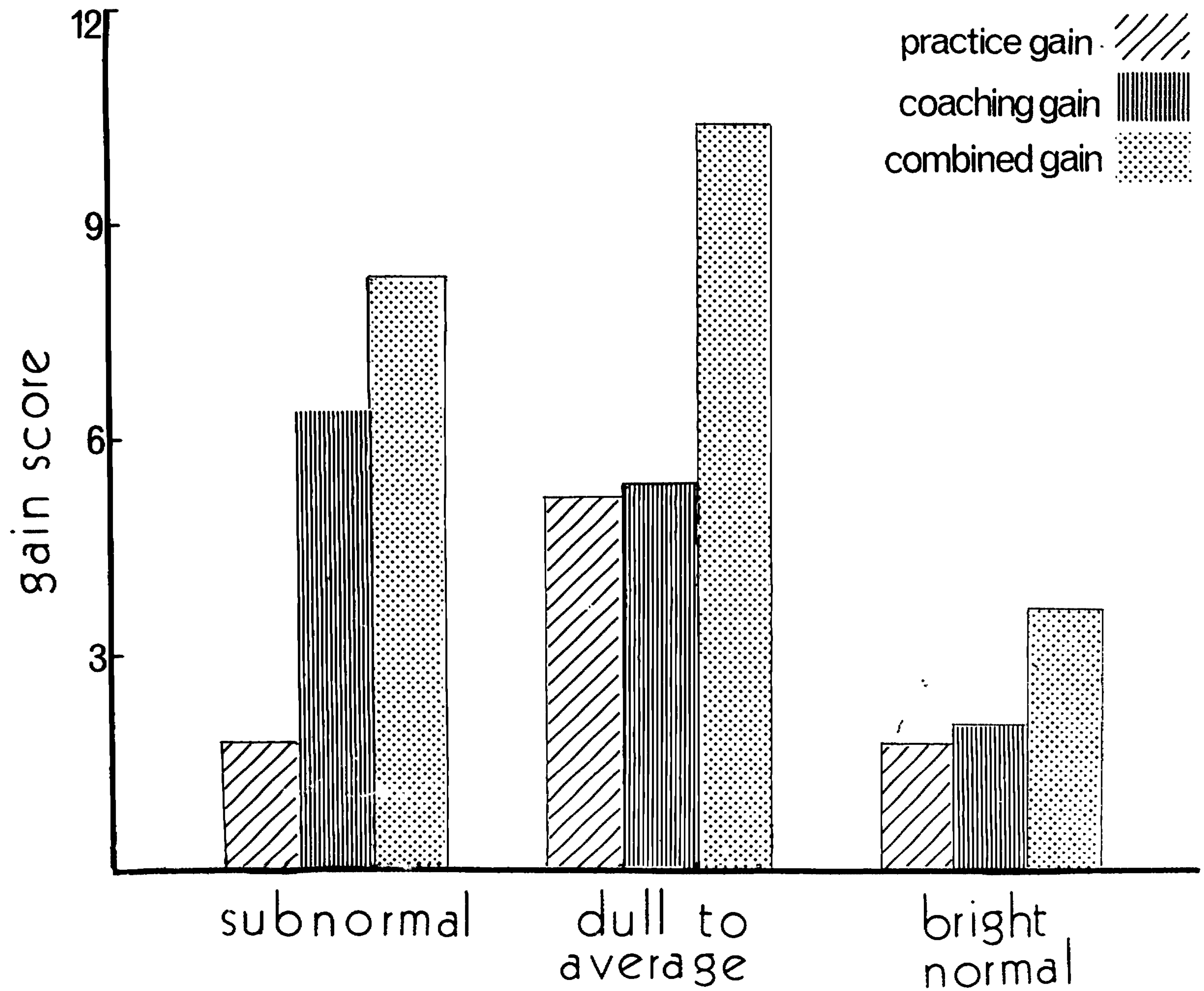
Table 7. Proportions of EMR subjects falling above the non-retarded (dull and bright) mean pretest scores in Study I and Study II

	Percent EMR subjects falling above the mean SLPT pretest score of the dull-to-average group		Percent EMR subjects falling above the mean SLPT pretest score of the bright normal group	
	EMR subjects in pretest	EMR subjects in posttest	EMR subjects in pretest	EMR subjects in posttest
Study I (special-class EMR subjects)	16%	36%	0%	13%
Study II (regular-class EMR subjects)	19%	61%	3%	36%

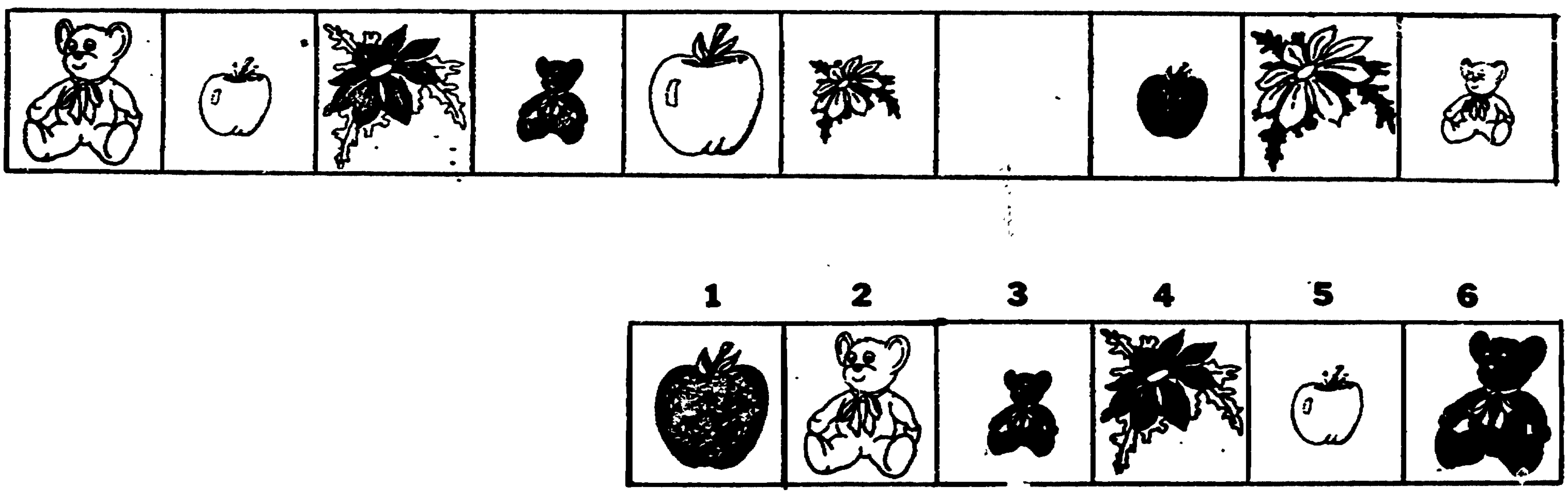
FIGURE CAPTIONS

Fig. 1. Sample item from the Series Learning Potential Test (SLPT, Babad and Budoff, 1971)

Fig. 2. Practice gain, coaching gain, and combined gain of the three groups.



practice gain
coaching gain
combined gain
gain score
12
9
6
3
subnormal
dull to average
bright normal


1
2
3
4
5
6