Chapter 13
Dynamic Assessment of Learning Potential

David Tzuriel

13.1 Introduction

The dynamic assessment (DA) of learning potential approach presented in this chapter is based mainly on Vygotsky's (1978) sociocultural theory, specifically the *zone of proximal development* concept, and Feuerstein's *mediated learning experience (MLE)* theory (Feuerstein et al. 1979) and Tzuriel's DA approach developed in the last three decades (Haywood and Tzuriel 1992; Tzuriel 1989, 1997, 2000, 2001, 2002; Tzuriel and Klein 1985). DA refers to an assessment, by an active teaching process, of a child's perception, learning, thinking, and problem solving. The process is aimed at modifying an individual's cognitive functioning and observing subsequent changes in learning and problem-solving patterns within the testing situation (Tzuriel 2001). The term *static* (or *standardized*) test refers to a test where the examiner presents items to the child and records his/her response without any attempt to intervene in order to change, guide, or improve the child's performance.

DA has been motivated by the inadequacy of conventional static tests to provide accurate information about the individual's learning ability, specific deficient functions, change processes, and mediation strategies that are responsible for cognitive modifiability. The need to develop DA tests has emerged because of criticism on static standardized tests and the difference in type of questions asked by DA as compared with standardized testing.

In the following sections of this chapter, I will discuss (a) the main criticism on standardized static tests, (b) the main goals of DA, (c) the major shifts of DA from standardized testing, (d) the major strategies of mediation in DA, (e) the use of DA in educational research, (f) the criticism of DA, and (g) why DA is not applied on a larger scale.

D. Tzuriel (✉)
School of Education, Bar Ilan University, Ramat Gan, Israel
e-mail: David.Tzuriel@biu.ac.il

M. Mo Ching Mok (ed.), *Self-directed Learning Oriented Assessments in the Asia-Pacific*, Education in the Asia-Pacific Region: Issues, Concerns and Prospects 18,
DOI 10.1007/978-94-007-4507-0_13,

13.2 Main Criticism on Standardized Static Tests

The major criticism against standardized testing can be summarized in the following main points:

(a) A frequent argument raised in the literature is that standardized static tests are biased toward minority groups and children with special needs and do not reflect their true ability. Children who come from low socioeconomic status (SES) families do not have adequate learning opportunities or efficient mediation from their parents and therefore fail in academic performance and/or in standardized tests. Their failure, however, does not reflect lack of intellectual abilities but rather lack of learning strategies, deficient cognitive functions (e.g., impulsivity), learning habits, self-efficacy in academic domains, and task-intrinsic motivation (Feuerstein et al. 1979).

(b) Another argument is that standardized tests are characterized many times by selective administration procedures and selective interpretation of results among high-risk children. For example, more lenient procedures (e.g., repeating instruction, showing more sympathy, allowing extra time, and giving hints) are used with children coming from high SES families than with children coming from low SES families. Although the test procedures are standardized, some examiners might use an "under-the-table" strategy of giving little cues for items not answered. This differential response might be on a subconscious or even a conscious level. In DA, on the other hand, mediation is "on-the-table" as the child is given "full-blooming" guidance and help. Another aspect of differential testing is selective interpretation of test results. Some examiners might judge a child's performance, especially a child with special needs, more strictly than a typically developing child and reach more severe conclusions than what actually is the child's level.

(c) A major argument against standardized tests is that motivational, emotional, and personality factors are not well taken. Research literature and teaching experience show that the motivational, emotional, and personality factors are no less important than the "pure" cognitive factors (Haywood 1968, 1971; Haywood and Lidz 2007; Tzuriel et al. 1988). Unfortunately, these factors are not given the proper attention in static tests or even totally neglected.

(d) The strongest argument against static tests is lack of information on learning and metacognitive processes. Those processes are of most importance in explaining the child's learning in the classroom and academic achievements. Teachers are interested in getting information on learning processes no less than on the relative standing of the child as compared with peers. As opposed to standardized tests, DA provides educators with data needed to suggest specific strategies for effective instruction and intervention. The different orientations of DA approach from static test approach derive from the major distinction in the type of questions asked by each approach. While in static testing the focus is on question of *what* is the level of the child's ability relative to same-age peers or what is the child's profile on certain subscales, in DA the questions are focused

on *how* the child *processes* the information, *what* are the specific cognitive functions responsible for the child's performance, *how* can we change and improve thinking and learning, and how cognitive, motivational, and emotional changes during testing can be used later to enhance the child's functioning in academic and nonacademic settings.

(e) Very frequently, static tests provide inadequate recommendations on remediation processes, specific interventions strategies, and prescriptive teaching. Many times, there is a "communication gap" between teachers and psychologists regarding translation of test findings into day-to-day teaching activities. It is common to find that teachers do not understand the terminology of static tests; the psychometric information is useless and barely translated to treatment strategies. Very frequently, psychologists do not have much experience with learning processes, and the static test data are not easily translated to specific recommendations.

13.3 Goals of DA

In order to understand deeply how DA is used and how it can help children with learning difficulties, we must understand the goals of DA. These goals may be summarized in the following:

(a) The first goal is to examine the capacity of the child to grasp the principle underlying an initial problem presented to the child and solve it correctly. This goal is very similar to the static test's goal, evaluating the manifested level of performance, or in Vygotsky's terms, the *actual level* of the *zone of proximal development*.

(b) The second goal is to assess the specific deficient cognitive functions as well as the adequate cognitive functions that are responsible for the child's failures and successes, respectively. Cognitive functions were defined as compounds of native abilities, learning habits, attitudes toward learning, motivational orientations, and cognitive strategies (Feuerstein et al. 1979). Adopting an information processing approach, Feuerstein suggested a list of deficient cognitive functions on the *input*, *elaboration*, and *output* phases of the mental act. For example, in the input phase one can identify difficulties in systematic exploratory behavior, simultaneous consideration of two or more sources of information, and spatial orientation. Deficient cognitive functions in the elaboration phase might be expressed by difficulties in planning behavior, comparative behavior, working memory, and episodic grasp of reality. Deficient cognitive functions in the output phase might be expressed by egocentric mode of communication, trial-and-error behavior, and projecting virtual relations. The deficient cognitive functions are considered as key elements for understanding children's performance. The modifiability of cognitive functions and operations (e.g., analogy, seriation) during DA is considered as an indicator for future changes, provided some treatment is given to modify them.

(c) The third goal of DA is to examine the nature and amount of investment required in order to teach the child a given principle or modify a deficient cognitive function. The examiner evaluates *how much* as well as *what types* of mediation are required in order to improve the child's cognitive functioning. This information is crucial in order to recommend later the type of mediation strategies needed as well as the required intensity.
(d) The fourth goal is to examine the extent to which the newly acquired principle is successfully applied in solving problems that become progressively more complex than the initial task. This goal is related to the level of internalization of learning and the amount of transfer the child's show in problem solving.
(e) The fifth goal is to examine the differential preference of the child for one or another modality of presentation of the problem (i.e., pictorial, linguistic, numerical). Understanding of the modality preference may help teachers in the future in designing intervention strategies and techniques.
(f) The sixth goal is to examine the differential effects of different training strategies given to the child to improve his/her functioning. It is important to understand what type of mediation is more effective especially in relation to the type of task that is given. The effects are measured by using the criteria of task level of novelty, level of complexity, language of presentation, and types of operation (i.e., analogy, syllogism, spatial orientation).

13.4 Major Shifts of DA from Standardized Testing

DA can be characterized by four major shifts from static standardized testing:

(a) *Goals of Testing.* The main goal in DA (see above goals of DA) is to assess learning potential and changes in task performance, cognitive functions, and nonintellective factors related to cognitive functioning. These changes are taken as indications for future changes, provided a cognitive intervention will be applied later to actualize the learning potential. In standardized testing, on the other hand, the main goal is to document the existing cognitive repertoire of the individual without any attempt to assess changes or learning processes.
(b) *Change in Nature of the Tasks.* Standardized tests are characterized by an emphasis on psychometric properties of the tasks, graduation of the difficulty levels of items, representation of children's capacities or knowledge, and administration procedures (e.g., test administration is terminated after several failures). Items are generally selected for the test if they coincide with psychometric properties (i.e., normal distribution, interitem reliability). In DA, on the other hand, the tasks are constructed on the basis of their "teaching potential"—the possibility of teaching important cognitive strategies, enhancing cognitive functions, and measuring cognitive changes. The items in DA are also graduated in terms of difficulty level, but the focus is on the teaching of cognitive strategies and operations so that learning of one task prepares the child to perform a more advanced task.

(c) *Change in Test Situation.* Since the objective of static tests is to compare an individual to his or her same-age peers, by definition, the test conditions require standardized stringent conditions for all examinees. Consequently, there is no room for teaching or interactive approach; examiners ask questions and examinees answer. Any guidance or help is perceived a transgression of the standardized conditions. Since the objective in DA is to change the individual's functioning within the test context, an active teaching approach is applied. Thus, there is a major shift in the role of the examiner from passive recording of the child's answers to an active mediation of cognitive strategies, rules, operations, and contents. In other words, while in standardized testing the examiner's roles are limited to administration of test items and later to scoring and interpretations, in DA the examiner intervenes to change the examinee's functioning and interprets future possible changes in view of current changes during assessment.

The DA interactive process is characterized by regulating the child's behavior through inhibition of impulsivity, sequencing and organizing the task dimensions, improving deficient cognitive functions, enriching the child's cognitive operations (e.g., comparative behavior, analogies, seriation) and task-related contents (e.g., labeling of relationships such as "opposite," "up–down"), and creating reflective and metacognitive processes.

The shift in test conditions might be symbolized by the sign frequently seen on the door of standardized testing rooms: "Silence! Testing in Progress." Contrary to the semiexperimental conditions required in standardized testing, in DA parents and teachers are often invited to observe the process. The observation may help later in explaining and reporting to parents the test results and in preparing for future cognitive intervention.

(d) *Change of Focus: From End Products to Process Orientation.* In standardized testing, the focus is on the end product of the mental act: the final answer. In DA, in contrast, the focus is on cognitive processes that bring about changes in specific deficient cognitive functions (e.g., impulsivity) and in nonintellective factors (e.g., need for mastery, resistance to mediation) that affect functioning. In other words, the emphasis is on process components, such as the nature of cognitive behavior, the learning process and strategies, and the specific interventions required to change them. While in standardized testing the emphasis is on the typical level of the child's performance, in DA the emphasis is on unique and qualitative aspects of the child's cognitive behavior. The questions asked in DA are "how" and "why" rather than "what" and "how much."

(e) *Change in Interpretation of Results*. While in standardized testing interpretation of results is based mainly on quantitative aspects, in DA it is based mainly on qualitative aspects of the child's performance, on analysis of the deficient cognitive functions, and on the mediational efforts required to modify them. The child's peak performance (i.e., independent performance after teaching) is taken as indicative of the child's ability rather than an average of all responses. Sometimes, only one bright answer provides a crucial indication of the child's learning potential, an indication that paves the way for deeper exploration of the possible factors that block the child from performing as well in other tasks.

Table 13.1 Major differences between DA and standardized testing

Dimensions	Dynamic assessment	Standardized testing
Goals of testing	Assessment of change	Evaluation of static performance
	Assessment of mediation	Comparison with peers
	Assessment of deficient cognitive functions	Prediction of future success
	Assessment of nonintellective factors	
Orientation	Processes of learning	End products (static)
	Metacognitive processes	Objective scores
	Understanding of mistakes	Profile of scores
Context of testing	Dynamic, open, and interactive	Standardized
	Guidance, help, and feedback	Structured
	Feelings of competence	Formal
	Parents and teachers can observe	Parents and teachers are not allowed to observe
Interpretation of results	Subjective (mainly)	Objective (mainly)
	Peak performance	Average performance
	Cognitive modifiability	
	Deficient cognitive functions	
	Response to mediation	
Nature of tasks	Constructed for learning	Based on psychometric properties
	Graduated for teaching	Termination after
	Guarantee for success	failures

These four major shifts from standardized testing to DA are summarized in Table 13.1.

13.5 Major Strategies of Mediation in DA

(a) *Improvement of (Deficient) Cognitive Functions.* The examiner should know how to identify the cognitive functions required for solution of a problem in the test and the mediation needed to improve the deficient cognitive functions.

(b) *Preparing the Child for Complex Tasks by Establishing Prerequired Thinking Behaviors.* Establishing prerequired thinking behaviors is carried out often by using mediation for transcendence and for self-regulation. Adequate initial investment in preparing the child brings about reduction of mediation efforts in later more abstract and complex problems. It is common to find children who solve difficult advanced problems much easier than the initial easy problems. Mediation of rules and principles (transcendence) has a motivational aspect as the child becomes independent of the examiner's mediation and enhances the child's sense of self-control. Mediation for self-regulation is carried out by focusing on systematic sequencing processes especially in complex problems

requiring an analytic approach. The examiner might ask the child to repeat the process of solution in order to crystallize the order of solution and to acquire feelings of mastery and efficiency.

(c) *Self-Regulation by Planning and Organization of the Solution.* One of the most frequent deficiencies among low-functioning children is impulsivity. Inhibition of impulsivity is done many times by decreasing the importance of time for performance. This is carried out by intentional delay of the child's response, longer exposure to the problem, systematic planning of the solution alternatives, verbalization of the problem, representation of the solution before pointing to the correct answer, and metacognitive analysis of the impulsive behavior. An efficient way of coping with impulsivity is by enriching the child's cognitive repertoire with thinking operations, comparative behavior, verbal tools, and hypothesis-testing techniques.

(d) *Enhancement of Reflective, Insightful, and Analytic Processes.* Enhancement of reflective, insightful, and analytic processes is carried out by focusing the child on the relation between his or her own thinking processes and the consequential cognitive performance. The focus is not on the end product but rather on the thinking process in the context of the required operations, type of task, and situation. Creation of insight is important for generalization and transfer of learning. It can be done by a dialogue with the child before solving the problem ("What should we look at before we will start to solve this problem?") or after the solution ("Why did you succeed in solving the problem that was so difficult for you to solve before?"). The most efficient way of enhancing reflective processes is by presenting the child with conflicts, incongruent information, intentional ambiguity, and absurd situations, which will bring about a need to close the cognitive gaps.

(e) *Teaching of Specific Contents that Are Related to the Task-Specific Context.* Teaching of specific contents (concepts, terms, relations) is not for the sake of language enrichment but for further use in problem-solving tasks. For example, the use of the terms up, down, vertical, horizontal, diagonal, similar, opposite, and different is necessary for performing the mental operation. The examiner can deviate for a short time from the task to teach and establish missing concepts and return later to the task to assess the performance efficiency and use of the newly acquired concepts.

(f) *Feedback on Success or Failure in the Learning Process.* The feedback given, which is one of the cornerstones in DA, is mutual—from the child and the examiner sides. It is especially important with low-performing children who are limited in their skills for giving feedback to themselves. This limitation is related to difficulties in self-correction and comparison of findings not only because of lack of knowledge and verbal tools of the children but also because of lack of orientation to make comparisons. Many tests are based on the assumption that trial-and-error behaviors will eventually bring the child to learn the correct answer. This assumption is wrong with regard to low-functioning children who are characterized by episodic grasp of reality. These children do not relate between their behavior and its consequences. A trial-and-error behavior

blocks their learning rather than facilitates it. The importance of feedback in DA derives from the examiner's ability to focus the child on the relation between behavior and consequence. The feedback is given not only on wrong answers but also on correct or partially correct answers, in order to teach self-correction. The goal of the feedback is beyond teaching the child a specific response. The aim is to teach insight, lawfulness, and meaning in relation to cognitive and emotional–motivational aspects.

(g) *Development of Basic Communication Skills and Adequate Response Style.* The mediation here is aimed at changing the child's response style so that problem solution will find a proper and efficient external expression. The examiner teaches the child how to communicate efficiently by the use of clear and accepted terms and avoiding egocentric communication. The examiner also teaches the child how to communicate precisely, justify the answer using logical arguments, and use verbal "codes" of expression and abstract high-order concepts rather than body gestures and facial expressions. It should be emphasized that previous communication style is not taken away before establishing new response styles.

13.6 Use of DA in Educational Research

The use of DA in educational research was aimed at (a) establishing the DA measures as more useful and accurate than standardized tests, especially with children showing learning difficulties and other clinical groups (Carlson and Wiedle 1992; Guthke and Stein 1996; Guthke and Wingenfeld 1992; Haywood and Lidz 2007; Hessels 2000; Resing 1997; Sternberg and Grigorenko 2002; Tzuriel 2001; Wiedl 2003), (b) validating theoretical concepts that are at the basis of DA (e.g., *zone of proximal development, structural cognitive modifiability)*, (c) demonstrating the effectiveness of DA in predicting school achievements, and (d) evaluating cognitive education programs. In the following sections, I will focus on two aspects: use of DA with children demonstrating learning difficulties and revealing the effectiveness of cognitive education programs by DA measures. For other aspects readers are referred to the respective literature (Haywood and Lidz 2007; Lidz and Elliott 2000; Sternberg and Grigorenko 2002; Tzuriel 2000, 2001).

13.6.1 DA with Children Demonstrating Learning Difficulties

DA was extensively used in research with children coming from low SES, minority ethnic groups, and different cultural backgrounds (Hessels 2000; Sternberg et al. 2002; Tzuriel and Kaufman 1999), as well as with children with learning and intellectual disability (Hessels-Schlatter 2002; Tzuriel 2000, 2001). In general, previous research has shown that standardized intelligence scores underestimate the cognitive

potential of children coming from low SES backgrounds, ethnic minority, and children with special needs, and that DA was proved to be more accurate in revealing their learning potential than static tests do (e.g., Guthke and Wingenfeld 1992; Hamers et al. 1991; Hessels 2000; Lidz and Elliott 2000; Resing 1997; Resing et al. 2009; Sternberg and Grigorenko 2002; Sternberg et al. 2002; Tzuriel 2000, 2001; Wiedl 2003).

DA results have been found to be more sensitive indicators of cognitive potential due to a variety of factors such as sociocultural deprivation, amount and quality of mediation provided at home, specific competencies for taking tests, interruptions in communication between examiner and examinee, test bias, and nonintellective factors such as self-confidence, need for mastery, and intrinsic motivation. By comparing static to DA measures, Guthke and Stein (1996) came to a conclusion that DA does not have a better predictive validity than static tests when used with typically developing students. However, in students with learning difficulties or atypical educational history, DA turned out to be a much better predictor of their future educational performance than static test scores. These findings support the conception of DA as an effective approach for revealing a "hidden" intellectual potential of special needs students. Sternberg et al. (2002) used DA with a group of rural Tanzanian school children ranging in grade levels from 2 to 5. The DA measures were largely based on fluid intellectual abilities such as syllogisms and sorting cards with different geometric figures. Children were assigned to experimental and control groups. The experimental group children received a short intervention phase for each test (well less than an hour per test) in which they were taught cognitive skills and strategies, whereas the control group children received no intervention. The findings showed significant pretest to posttest improvement across different tests in the experimental group as compared with the control group. Furthermore, posttest scores on the dynamic tests (administered in the experimental group only) were better predictors of reference ability and achievement measures than were pretest scores. One of the conclusions of this study, as expected, is that children growing up in difficult circumstances seem to have important intellectual abilities not measured by static tests.

In one of the earlier studies with young children, Tzuriel and Klein (1985) administered the *Children's Analogical Thinking Modifiability* (CATM) test to four groups of children: disadvantaged and advantaged kindergarten children, kindergarten children identified with special needs, and older intellectually disabled (ID) children with mental age equal to kindergarten level. The CATM is composed of three sets of analogies given in preteaching, teaching, and postteaching phases. The operation of analogy has been considered by many authors as a powerful tool for a wide range of cognitive processes and as a principal operation for problem-solving activities (Goswami 1991; Holyoak and Thagard 1997; Gentner and Markman 1997).

The CATM test is composed of 14 items for each phase of administration (preteaching, teaching, and postteaching) and 18 colored blocks that are used to present and solve the analogies. The CATM items, graduated in level of difficulty, require a relatively higher level of abstraction and various cognitive functions. Examples of items from the CATM test are portrayed in Fig. 13.1.

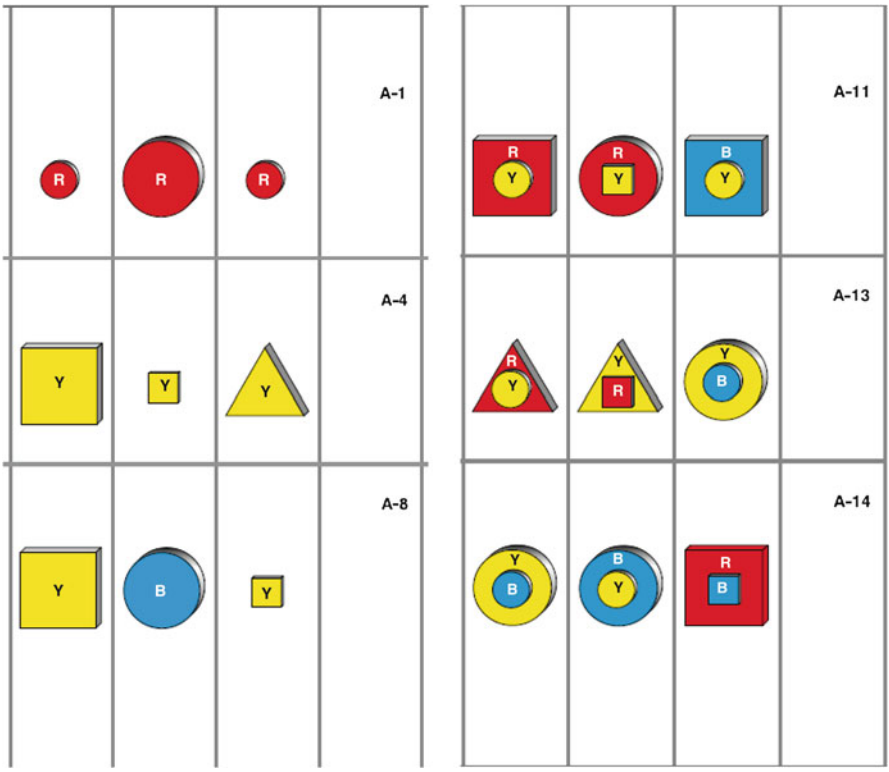

Fig. 13.1 Examples of items from the Children's Analogical Thinking Modifiability (CATM) test (R = Red, B = Blue, Y = Yellow)

In item 13, for example (see Fig. 13.1), the child has to compare the relations of colors in the first pair of the problem, find the rules of the relations, and apply them in the second pair. In the first pair, the relation of colors is opposite: *top*-yellow changes to *bottom*-yellow and *bottom*-red changes to *top*-red. If the rule of opposite is applied in the second pair, then the *top*-blue changes to *bottom*-blue and *bottom*-yellow changes to *top*-yellow. After finding the correct colors, the child can analyze the relations for the other two dimensions of shape and size (of both top and bottom components).

During the teaching phase, the child is mediated to (a) search for relevant dimensions required for the analogical solution, (b) understand transformational rules and analogical principles, (c) search systematically for correct blocks, and (d) improve efficiency of performance.

The CATM may be scored by two methods: "all-or-none" (e.g., a score of 1 is given to full answer) or "partial credit" (e.g., a score of 1 is given for each correct dimension of color, shape, and size). The findings showed that the highest gains from pre- to postteaching phases of the CATM test were found among disadvantaged and advantaged children as compared with children with needs for special

education and ID children, who showed small gains. The ID group, however, showed significant improvement when a "partial credit" scoring method was applied. This last finding indicates that the ID group had difficulty in integration of all sources of information and therefore showed modifiability only according to the "partial credit" method. Higher levels of functioning were found for all groups on the CATM than on a static test, the Raven's Colored Progressive Matrices (RCPM Raven 1956). The differences were especially articulated when the analogical items of the RCPM were compared to the analogical problems of the CATM. For example, the advantaged and disadvantaged children scored 69% and 64% on the CATM, respectively, as compared to 39% and 44% on the RCPM, respectively.

In another study on children with special needs, Tzuriel and Caspi (1992) compared deaf children with hearing children on both DA and standardized measures. The kindergarten deaf children were matched to hearing children on variables of age, sex, and a developmental visual-motor test. Both groups were tested on the CATM and RCPM tests. The findings showed that on the CATM-postteaching phase, the hearing and deaf children scored 66% and 54% ("all-or-none" scoring method) and 86% and 81% ("partial credit" scoring method), respectively, as compared to 42% and 39% on the RCPM, respectively. These findings indicate that both groups have a higher level of learning potential than is indicated by static test scores. Comparison of pre- to postteaching tests revealed that the deaf children performed lower than the hearing children on the preteaching test but showed greater improvement after the teaching phase; no significant group differences were found in the postteaching test.

Previous studies with minority and culturally different children have shown that DA provides information different from conventional static tests. Guthke and Al-Zoubi (1987) compared a sample of 200 grade 1 children in Germany to a comparable Syrian sample on both a static measure—the Colored Progressive Matrices (CPM)—and a DA measure. The findings showed that the German children scored significantly higher than did the Syrian children. However, after a training phase, there was only a slight difference between the two groups. These results were interpreted as an indication that both ethnic groups have the same intellectual endowments. Similarly, Hessels and Hamers (1993) reported that although minority children scored significantly lower than Dutch children on learning potential tests, the differences were markedly smaller than with IQ tests. In South Africa, Skuy and Shmukler (1987) and Shochet (1992) used the *Learning Potential Assessment Device* (Feuerstein et al. 1979) and other psychometric tests with groups of children and students of Indian, Black, and "colored" origin. Skuy and Shmukler (1987) reported that although mediation was not generally effective in producing change on transfer measures, it was effective with a subgroup of colored high academic status students. The group that benefited most from mediation was the high academic status colored students. Shochet (1992) investigated the predictability of success in the first year of studies in the university using indexes of cognitive modifiability taken before admission on a disadvantaged student population. The findings showed significant prediction among "less modifiable" students but not among the "more modifiable" students (modifiability was measured by DA prior to start of the studies). It was surmised

that they are less susceptible to being modified during the first year, either by direct exposure or by mediated learning experience (MLE).

A unique cross-cultural study was carried out by Tzuriel and Kaufman (1998) on a group of newly arrived Ethiopian children, in grade 1, who immigrated to Israel in the 1990s. They were compared with grade 1 Israeli-born children using static and DA tests. A central question that has been raised recently with new Ethiopian immigrants to Israel is how to assess their learning potential, especially in view of the inadequacy of standard testing procedures to reflect this population's cognitive functioning accurately. The question, however, transcends the specific context of the Ethiopian Jews. Theoretically, it is related to issues such as the influence of cultural changes on the individual's cognitive functioning, internalization of novel symbolic mental tools with transition from one culture to another, and resilience in coping with cultural incongruences. Pragmatically, this question applies to a variety of populations who, for sociohistorical reasons, live as subcultures within a broad culture and whose members might be penalized by inadequate diagnostic procedures.

It should be noted that the Ethiopian immigrants, upon arrival to Israel, had to overcome a gap of civilization and information of many years and had to adapt to the Israeli society. Coming from an illiterate society where their rich culture was transmitted orally, they had to go, upon arrival to Israel, through rapid change and adjust to differences in both material and symbolic tools. All children were administered the *Raven's Colored Progressive Matrices* (CPM Raven 1956), the CATM test, and the *Children's Inferential Thinking Modifiability* test (CITM Tzuriel 1989); the last two are DA measures. The CITM test, which is presented using verbal and pictorial modalities, taps several cognitive functions such as comparative behavior, systematic exploratory behavior, self-regulation of impulsivity, and inferential-hypothetical operations. An example of an item from the CITM is presented in Fig. 13.2.

The CITM test is composed of sets of problems for preteaching, teaching, postteaching, and transfer phases. After presentation of a set of 24 familiar pictures (e.g., clothes, animals, furniture) and naming them, the child is given two example problems and is instructed in the rules and procedures for solving them. Each problem consists of rows of figures, each row presenting partial information about the possible location of objects in houses with different colored roofs. The child is required to compare the information presented in the rows, infer the exact location of the objects, and place them in their right houses. The basic rule is that pictures on the left should be in houses with lines on the right. In Fig. 13.2, for example, the bicycle and cabinet in row 1 should go to the black and red houses, but we do not know which picture goes to which house. The child has to compare rows 1 and 2, identify the common elements, and make the inference (e.g., "the bicycle and the black house appear in both rows therefore the bicycle goes to the black house at the top of the page").

The CITM requires planning behavior, systematic exploratory behavior, a strategic and analytic approach, need for accuracy, and control of impulsivity. Although the tasks were novel to the children in both groups, the mental operations required to solve them are relatively familiar and to some degree are also practiced among the Israeli-born

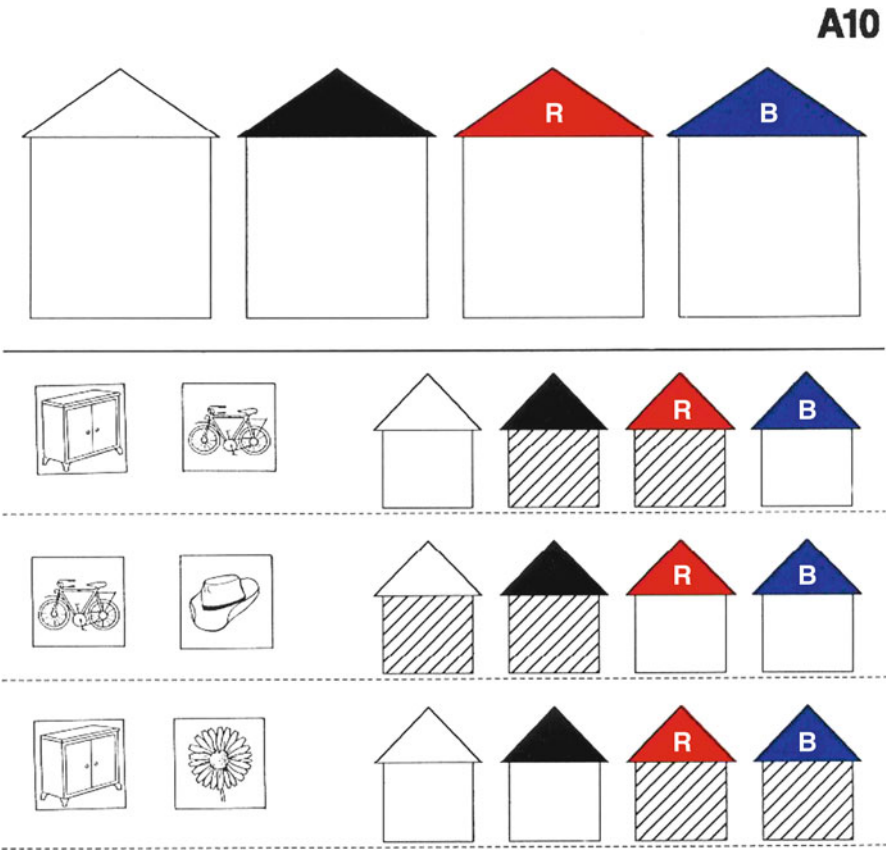

Fig. 13.2 Example of an item from the Children's Inferential Thinking Modifiability (CITM) test (R = Red, B = Blue)

children. For the Ethiopian children, however, these mental activities are new and have no similarity to the type of activities practiced or transmitted in their culture.

The findings showed clearly that the Israeli-born group scored higher than the Ethiopian group on the CPM (static) and the preteaching DA tests. However, the improvement from pre- to postteaching phases of the DA was higher for the Ethiopian than for the Israeli-born group. The findings on The CITM are presented in Fig. 13.3.

As can be seen in Fig. 13.3, the Ethiopian children narrowed the gap on the postteaching phase of the CITM; differences on both postteaching and transfer problems were not significant. The lack of significant differences on the transfer items indicates that the Ethiopian children could benefit from the mediation given to them, internalize the rules, and use them efficiently in the transfer items. The large cognitive change among the Ethiopian children supports both Vygotsky's (1978) ZPD and Feuerstein et al.'s (1979) cognitive modifiability constructs.

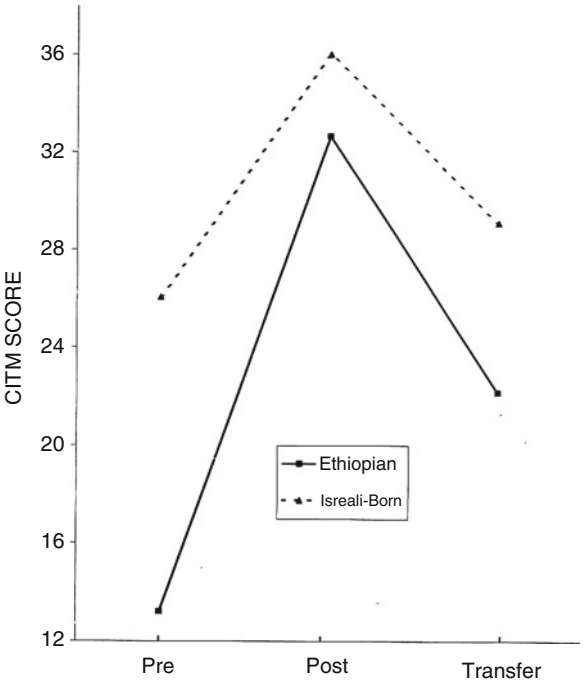

Fig. 13.3 CITM test preteaching, postteaching, and transfer scores of Israeli-born and Ethiopian children (Copied by permission from the *Journal of Cross-Cultural Psychology*)

One of the most intriguing and impressive findings was on the classification phase of the CITM. After finishing the inferential task, children are asked to classify the pictures (cards) presented during the earlier section to categories. There are six categories (e.g., animals, cloths, figures, furniture, means of transportation, plants); each category contains four pictures. Each correctly solved category can get a score of 2 and a maximal score of 12 for all categories. After the first classification phase, all children received a simple mediation phase that lasted between 1 and 2 min in which the principle of classification was explained.

The Ethiopian children achieved a dramatic and significant gain from .70 to 9.00 as compared with a gain from 10.20 to 12.00 among the Israeli-born children who reached a ceiling. It should be noted that the low initial score of the Ethiopian children was not a result of inadequate instruction but of a different understanding of what is expected to perform. For example, a typical classification of objects in the premediation phase among Ethiopian children could be a donkey, a leaf, and a circle. When asked why these three pictures are classified together, the answer was "because the donkey eats the leaf by the well (circle)." After a simple explanation of the meaning of a class (e.g., donkey, dog, cat, and bird; all of them belong to the family of animals), the improvement was drastic. These results coincide with cross-cultural research findings indicating that individuals in many non-Western nations classify items into functional rather than into taxonomic categories (e.g., Greenfield 1997).

In a recent study, ethnic minority children in the Netherlands were compared to indigenous children on a DA test: the Seria-Think Instrument (Tzuriel 2000) using a graduated prompt technique (Resing et al. 2009). The findings showed that children tested by DA changed their strategy behavior into the direction of a more advanced strategy and that this change was the largest for the initial weaker scoring ethnic minority children. More specifically, ethnic minority children initially needed more, but then progressively needed fewer, cognitive hints than did the indigenous children. These findings show that ethnic minority children need support in order to know what to solve and how to do it. Once the situation was clarified, they showed greater progression toward superior strategy use.

13.6.2 Evaluating the Effects of Cognitive Education Programs by DA

DA has been used frequently to assess the effectiveness of cognitive intervention programs. The rationale of using DA is matching the declared objective of the cognitive program (e.g., "learning how to learn") with criterion measures of change and modifiability. DA has been used for evaluating four cognitive intervention programs: *Instrumental Enrichment* (*IE,* Feuerstein et al. 1980), *Bright Start* (Haywood et al. 1986), *Peer Mediation for Young Children (PMYC,* Shamir and Tzuriel 2004; Tzuriel and Shamir 2007, 2010), and the *Analogical Reasoning Program (ARP,* Tzuriel and George 2009). The findings of several studies show clearly that the effectiveness of the program could be revealed only when DA approach was applied. Because of space limitation, I will present here two recent studies, one on *Bright Start* and the second on PMYC program. For a detailed review on revealing the effectiveness of DA in evaluating cognitive education programs, see Tzuriel (2011).

In the first study (Tzuriel et al. 1999), a sample of kindergartners received the Bright Start in their classrooms ($n=82$) and was compared to a group of children ($n=52$) who received a basic skills program. The Bright Start program was applied for 10 months, during which the children in the experimental group received five of the seven small-group units: self-regulation, quantitative relations (number concepts), comparison, classification, and role-taking. The small-group lessons were taught three times a week, each session for a period of 20 min, for a total of 1 h per week and a total number of 32 h for the academic year. The comparison group was given the basic skills program during the academic year, and the teachers were visited periodically to observe their skills-based program. Two DA instruments were administered: the CATM and a young children's version of the *Complex Figure* test (Tzuriel and Eiboshitz 1992). Since the finding of the Complex Figure is very similar to those of the CATM, only the CATM findings are reported here.

After gathering the preintervention data, we realized that the cognitive scores of the experimental group were lower than those of the comparison group. Unfortunately, there was no possibility of random assignment of children in each class to the treatment groups without raising the parents' resentment. It would also have been confusing to the kindergarten teacher who would have had to implement both

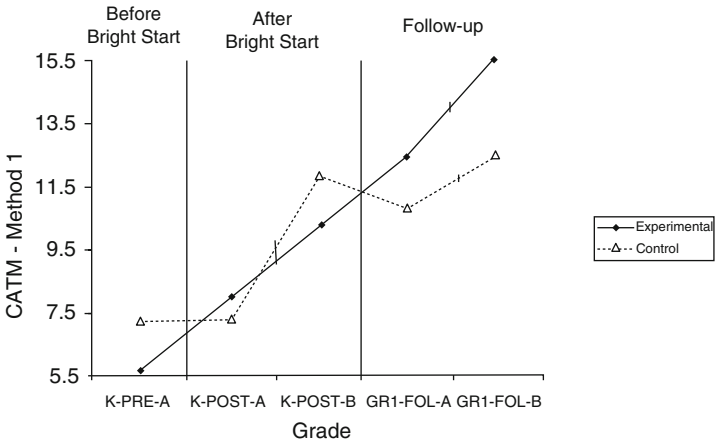

Fig. 13.4 The CATM Scores of the Experimental and Comparison Groups Before and After the Bright Start Program, and in the Follow-up Phase (Copied by permission from *Early Childhood Research Quarterly, ECRQ*) [(K = Kindergarten, GR1 = Grade 1, Pre-A = Pre-Intervention Preteaching Problems (A), Post-A = Post-Intervention Preteaching problems (A), Post-B = Post-Intervention Postteaching Problems (B), Fol-A = Follow-up Preteaching problems (A), Fol-B = Follow-up Postteaching Problems (B)]

programs within one class. We had to rely, therefore, on supervisors' assessments of children's background as a basis for equating the treatment groups. This eventually proved to be not completely accurate.

Group comparison on CATM and Complex Figure pre- and postteaching scores was carried out at the end of the intervention and in a follow-up phase 1 year after the end of the program. A MANOVA of treatment (experimental vs. comparison) by phase (pre- vs. postteaching) and by grade (*K* vs. grade 1) was carried out on the CATM scores. The analysis revealed a significant triadic interaction of treatment by grade by pre-/postteaching, $F_{(2, 69)} = 4.27$, $p < .02$. The interaction is portrayed in Fig. 13.3. For comparative reasons, the CATM scores at the start of the program are also plotted in Fig. 13.3; however, the analysis is based only on students who participated in the follow-up.

Figure 13.4 shows both static and DA results. The static tests results are portrayed in CATM scores before and after the intervention (set A, preteaching). The findings show that children in the experimental group made higher improvement on the CATM scores (set A) from preintervention (*K*-Pre-A) to postintervention (*K*-Post-A) phase. When the CATM was administered in a DA procedure, the findings were intriguing. While at the end of the program (*K*) the comparison children improved their performance from the pre- to postteaching phase of the DA test more than did the experimental children, in the follow-up year (grade 1) the trend was reversed! The experimental group showed higher improvement from pre- to postteaching than did the comparison group.

These results in grade 1 were interpreted as an indication for a "snowball" effect of the "learning to learn" treatment. According to the "snowball effect," treatment

effects gain power with time without any additional treatment, which is to be expected when the treatment is designed to enhance "learning to learn" skills. Further support for the "snowball effect" was found when cognitive modifiability indices were taken as the dependent variable. Cognitive modifiability indices were calculated by regression analysis in which the residual postteaching scores were derived after controlling for the preteaching score (see Embreston 1987, 1992).

A MANOVA of treatment by grade (2×2) applied on the CATM cognitive modifiability indices revealed a significant overall interaction of treatment by grade, $F_{(2, 69)} = 10.08$, $p < .0001$. This finding indicates higher improvement of the cognitive modifiability scores in the experimental than in the comparison group, from kindergarten to first grade.

13.7 Criticism on Dynamic Assessment

A frequent criticism mentioned in the literature is that DA takes more time to administer and requires more skill, better training, more experience, and greater effort than static testing do (Frisby and Braden 1992). It is true that the professional skill necessary to do DA effectively is not currently taught in typical graduate psychology programs, so practitioners must be trained in intensive workshops long after they have been indoctrinated in the "laws" of static, normative testing (Haywood and Tzuriel 2002). Even with excellent training, DA examiners must exercise considerable subjective judgment in determining (a) what cognitive functions are deficient and require mediation, (b) what kinds of mediation to dispense, (c) when further mediation is not needed, and (d) how to interpret the difference between premediation and postmediation performance. It seems somehow disingenuous to complain that DA requires special knowledge and special skills when its benefits are directly related to such knowledge and skills and in turn have benefits for the children.

Another criticism is that the extent to which *cognitive modifiability* is generalized across domains (i.e., analogical, numerical) needs further investigation. Related to this criticism is the question of how to translate the DA findings into effective instruction and intervention. This aspect is considered as a major educational advantage over static testing.

The relative lack of reliability is another major criticism. Establishing reliability and validity of DA is much more complex than validation of static testing because it has a broader scope of goals. The question of reliability is a pressing one, especially so given that one sets out deliberately to change the very characteristics that are being assessed. At least a partial solution is to insist on very high reliability of the tasks used in DA when they are given in a static mode, i.e., without interpolated mediation. Another solution is to use interjudge reliability based on observations. This aspect has been studied to some extent (e.g., Tzuriel and Samuels 2000) but not yet sufficiently.

Another persistent problem is how to establish the validity of DA. Ideally, one would use both static testing and DA with one group of children and static, normative

ability tests with another group. The essential requirement would be that a subgroup of the DA children would have to be given educational experiences that reflected the within-test mediation that helped them to achieve higher performance in DA. The expectation would be that static tests would predict quite well the school achievement of both the static testing group and that subsample of the DA group that did not get cognitive educational follow-up. Static tests should predict less well the achievement of the DA-cognitive education group; in fact, the negative predictions made for that group should be defeated to a significant degree (Haywood and Tzuriel 2002).

One of the criticism raised by Frisby and Braden (1992) is that the literature is replete with evidence showing a strong relation between IQ and school achievement (r=.71). The question therefore is why applying a DA approach if so much of the variance in school learning is explained by standardized testing? The last point means that nearly 50% of the variance in learning outcomes for students can be explained by differences in psychometric IQ. My answer to the last point, being loyal to a meditational approach of inquiring and probing, is by asking three extremely important questions (Tzuriel 1992). These questions are graduated from light to heavy:

(a) What causes the other 50% of achievement variance?
(b) When IQ predicts low achievement, what is necessary to defeat that prediction?
(c) What factors influencing the unexplained variance can help to defeat the prediction in the explained variance?

13.8 Why DA Is Not Applied on a Larger Scale?

One might well ask why, if DA is so rich and rewarding, it is not more widely applied? Here are some possible answers (Karpov and Tzuriel 2009):

- One apparent reason is that it is not taught in graduate school yet.
- School psychologists often have "client quotas" to fill, and DA is far more time-consuming that is static testing, so their supervisors do not permit it.
- The school personnel who ultimately receive the psychologists' reports typically do not expect DA and do not yet know how to interpret the data or the recommendations, and psychologists have not been good enough about helping them on that score.
- There is a certain inertia inherent in our satisfaction with being able to do what we already know how to do and to do it exceptionally well. Even so, as we have observed before, "what is not worth doing is not worth doing well!"

The question of what should be done is complex as the answer depends on a myriad of interrelated factors. Haywood (2008) suggested that the most urgent task is to explore and incorporate new models of the nature of human ability. He suggested, as one such model, a "transactional" perspective on human ability with three major

dimensions: intelligence, cognitive processes, and motivation, especially task-intrinsic motivation. The concept of intelligence, then, is not seen as useless or as antithetical to the notion of cognitive processes, structures, or strategies but as a construct that does not explain all that we know about individual differences in learning and performance effectiveness. We can supplement its explanatory value by adding the dimensions of cognitive processes and motivation. One should proceed from some such model of the nature of ability to define what it is that we wish to assess and only then to construct instrument for assessing individual differences in that set of variables.

References

Carlson, J. S., & Wiedle, K. H. (1992). Principles of dynamic assessment: The application of a specific model. *Learning and Individual Differences, 4*, 153–166.

Embretson, S. E. (1987). Toward a development of a psychometric approach. In C. S. Lidz (Ed.) *Dynamic assessment* (pp. 141–170). New York: Guilford Press.

Embreston, S. E. (1992). Measuring and validating cognitive modifiability as ability: A study in the spatial domain. *Journal of Educational Measurement, 29*, 25–50.

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

Feuerstein, R., Rand, Y., Hoffman, M. B., & Miller, R. (1980). *Instrumental enrichment*. Baltimore: University Park Press.

Frisby, C. L., & Braden, J. P. (1992). Feuerstein's dynamic assessment approach: A semantic, logical and empirical critique. *Journal of Special Education, 26*, 281–301.

Gentner, D., & Markman, A. B. (1997). Structure mapping in analogy and similarity. *American Psychologist, 52*, 45–56.

Goswami, U. (1991). Analogical reasoning: What develops? A review of research and theory. *Child Development, 62*, 1–22.

Greenfield, P. M. (1997). You can't take it with you: Why ability assessments don't cross cultures. *American Psychologist, 52*, 1115–1124.

Guthke, J., & Al-Zoubi, A. (1987). Kulturspezifische differenzen in den Colored Progressive Matrices (CPM) und in einer Lerntestvariante der CPM [Culture specific differences in the Colored Progressive Matrices (CPM) and in learning potential version of the CPM]. *Psychologie in Erziehung und Unterricht, 34*, 306–311.

Guthke, J., & Stein, H. (1996). Are learning tests the better version of intelligence tests? *European Journal of Psychological Assessment, 12*, 1–13.

Guthke, J., & Wingenfeld, S. (1992). The learning test concept: Origins, state of the art, and trends. In H. C. Haywood & D. Tzuriel (Eds.), *Interactive assessment* (pp. 64–94). New York: Springer.

Hamers, J. H. M., Hessels, M. G. P., & Van Luit, J. E. H. (1991). *Learning potential test for ethnic minorities: Manual and test*. Lisse: Swets & Zeitlinger.

Haywood, H. C. (1968). Motivational orientation of overachieving and underachieving elementary school children. *Journal of Personality, 30*, 63–74.

Haywood, H. C. (1971). Individual differences in motivational orientation: A trait approach. In P. I. Day, D. E. Berlyne, & D. E. Hunt (Eds.), *Intrinsic motivation: A new direction in education* (pp. 113–127). New York: Holt, Rinehart, & Winston.

Haywood, H. C. (2008). Twenty years of IACEP, and a focus on dynamic assessment: Progress, problems, and prospects. *Journal of Cognitive Education and Psychology, 7*, 419–442.

Haywood, H. C., & Lidz, C. S. (2007). *Dynamic assessment in practice: Clinical and educational applications*. New York: Cambridge University Press.

Haywood, H. C., & Tzuriel, D. (Eds.). (1992). *Interactive assessment*. New York: Springer.
Haywood, H. C., & Tzuriel, D. (2002). Applications and challenges in dynamic assessment. *Peabody Journal of Education, 77*, 38–61.
Haywood, H. C., Brooks, P. H., & Burns, M. S. (1986). Stimulating cognitive development at developmental level: A tested, non-remedial preschool curriculum for preschoolers and older retarded children. In M. Schwebel & C. A. Maher (Eds.), *Facilitating cognitive development: Principles, practices, and programs* (pp. 127–147). New York: Haworth Press.
Hessels, M. G. P. (2000). The learning potential test for ethnic minorities: A tool for standardized assessment of children in kindergarten and the first years of primary school. In C. S. Lidz & J. G. Elliott (Eds.), *Dynamic assessment: Prevailing models and applications* (pp. 109–131). New York: Elsevier.
Hessels, M. G. P., & Hamers, J. H. M. (1993). A learning potential test for ethnic minorities. In J. H. M. Hamers, K. Sijtsma, & A. J. J. M. Ruijssenaars (Eds.), *Learning potential assessment* (pp. 285–311). Lisse: Swets & Zeitlinger.
Hessels-Schlatter, C. (2002). Dynamic test to assess learning capacity in people with severe impairments. *American Journal on Mental Retardation, 107*(5), 340–351.
Holyoak, K. J., & Thagard, P. (1997). The analogical mind. *American Psychologist, 52*, 35–44.
Karpov, Y. & Tzuriel, D. (2009). Dynamic assessment: Progress, problems, and prospects. *Journal of Cognitive Education and Psychology, 8*, 228–237.
Lidz, C. S., & Elliott, J. G. (Eds.). (2000). Advances in cognition and educational practice. In *Dynamic assessment: Prevailing models and applications*. Oxford: Elsevier.
Raven, J. C. (1956). *Guide to using the Colored Progressive Matrices, sets A, Ab, and B*. London: Lewis.
Resing, W. C. M. (1997). Learning potential assessment: The alternative for measuring intelligence? *Educational and Child Psychology, 14*, 68–82.
Resing, W. C. M., Tunteler, E., de Jong, F. M., & Bosma, T. (2009). Dynamic testing in indigenous and ethnic minority children. *Learning and Individual Differences, 19*, 445–450.
Shamir, A., & Tzuriel, D. (2004). Children's mediational teaching style as a function of intervention for cross-age peer-mediation. *School Psychology International, 25*, 58–97.
Shochet, I. M. (1992). A dynamic assessment for undergraduate admission: The inverse relationship between modifiability and predictability. In H. C. Haywood & D. Tzuriel (Eds.), *Interactive assessment* (pp. 332–355). New York: Springer.
Skuy, M., & Shmukler, D. (1987). Effectiveness of the learning potential assessment device for Indian and "colored" South Africans. *International Journal of Special Education, 2*, 131–149.
Sternberg, R. J., & Grigorenko, E. L. (2002). *Dynamic testing: The nature and measurement of learning potential*. New York: Cambridge University Press.
Sternberg, R. J., Grigorenko, E. L., Ngorosho, D., Tantufuye, E., Mbise, A., Nokes, C., Jukes, M., & Bundy, D. A. (2002). Assessing intellectual potential in rural Tanzanian school children. *Intelligence, 30*, 141–162.
Tzuriel, D. (1989). Inferential cognitive modifiability in young socially Israeli-born and advantaged children. *International Journal of Dynamic Assessment and Instruction, 1*, 65–80.
Tzuriel, D. (1992). The dynamic assessment approach: A reply to Frisby and Braden. *Journal of Special Education, 26*, 302–324.
Tzuriel, D. (1997). A novel dynamic assessment approach for young children: Major dimensions and current research. *Educational and Child Psychology, 14*, 83–102.
Tzuriel, D. (2000). Dynamic assessment of young children: Educational and intervention perspectives. *Educational Psychology Review, 12*, 385–435.
Tzuriel, D. (2001). *Dynamic assessment of young children*. New York: Kluwer Academic/Plenum Press.
Tzuriel, D. (2002). Dynamic assessment of learning potential. In J.W. Guthrie (Ed.), *Encyclopedia of education* (2nd ed., pp. 127–131). New York: McMillan Press.
Tzuriel, D. (2011). Revealing the effects of cognitive education programs by dynamic assessment. *Assessment in Education: Principles, Policy and Practice, 18*(2), 113–131.

Tzuriel, D., & Caspi, N. (1992). Dynamic assessment of cognitive modifiability in deaf and hearing preschool children. *Journal of Special Education, 18*, 113–131.
Tzuriel, D., & Eiboshitz, Y. (1992). A structured program for visual motor integration (SP-VMI) for preschool children. *Learning and Individual Differences, 4*, 103–124.
Tzuriel, D., & George, T. (2009). Improvement of analogical reasoning and academic achievements by the Analogical Reasoning Program (ARP). *Educational and Child Psychology, 29*, 71–93.
Tzuriel, D., & Kaufman, R. (1999). Mediated learning and cognitive modifiability: Dynamic assessment of young Ethiopian immigrants in Israel. *Journal of Cross-Cultural Psychology, 30*, 359–380.
Tzuriel, D., & Kaufman, R. (1998). Mediated learning and cognitive modifiability: Dynamic assessment of young Ethiopian immigrants in Israel. *Journal of Cross-Cultural Psychology, 13*, 539–552.
Tzuriel, D., & Klein, P. S. (1985). Analogical thinking modifiability in disadvantaged, regular, special education and mentally retarded children. *Journal of Abnormal Child Psychology, 13*, 539–552.
Tzuriel, D., & Samuels, M. T. (2000). Dynamic assessment of learning potential: Inter-rater reliability of deficient cognitive functions, type of mediation, and non-intellective factors. *Journal of Cognitive Education and Psychology, 1*, 41–64.
Tzuriel, D., & Shamir, A. (2007). The effects of peer mediation with young children (PMYC) on children's cognitive modifiability. *British Journal of Educational Psychology, 77*, 143–165.
Tzuriel, D., & Shamir, A. (2010). Mediation strategies and cognitive modifiability in young children as a function of peer mediation with young children (PMYC) program and training in analogies versus math tasks. *Journal of Cognitive Psychology and Education, 9*, 48–72.
Tzuriel, D., Samuels, M. T., & Feuerstein, R. (1988). Non-intellective factors in dynamic assessment. In R. M. Gupta & P. Coxhead (Eds.), *Cultural diversity and learning efficiency: Recent developments in assessment* (pp. 141–163). London: NFER-Nelson.
Tzuriel, D., Kaniel, S., Kanner, E., & Haywood, H. C. (1999). The effectiveness of Bright Start program in kindergarten on transfer abilities and academic achievements. *Early Childhood Research Quarterly, 114*, 111–141.
Vygotsky, L. S. (1978). *Mind in society*. Cambridge, MA: Harvard University Press.
Wiedl, K. H. (2003). Dynamic testing: A comprehensive model and current fields of application. *Journal of Cognitive Education and Psychology, 3*, 93–119.