

**14th Edition of the International Conference on Sciences
of Education. Learning for Life - ICSED 2016, 12-13
May 2016, Suceava (Romania) & Chernivtsi (Ukraine)**

**Studies and Current Trends in Science of Education
Editors Otilia CLIPA & Constantin-Florin DOMUNCO
ISBN: 978-973-166-435-4; e-ISBN: 978-973-166-475-0**

New Orientations in the Evaluation of Psychical Development and Academic Achievements at Early Ages

Maria Magdalena STAN

pp.311-322

©2016 The Authors & LUMEN Publishing House.

Selection, peer review and publishing under the responsibility of the editors.

How to cite: Stan, M. M. (2016). New Orientations in the Evaluation of Psychical Development and Academic Achievements at Early Ages. In O. Clipa & C. F. Domunco, (eds.), Studies and Current Trends in Science of Education (pp. 311-322). Iasi, Romania: Editura LUMEN.

New Orientations in the Evaluation of Psychological Development and Academic Achievements at Early Ages

Maria Magdalena STAN¹

Abstract

The assumptions of the present paper represent the methodological and practical necessity of an objective and unitary evaluation at the academic studied level, mainly at early ages (the task being the first literacy – the learning of writing - reading and ciphering) with a view to monitorize, to diagnose and prognosticate the academic results and to certify the level of academic achievement.

The present paper proposes a modality of evaluating academic achievement as well as the development level at early ages.

Keywords

Evaluation of academic achievement, psychological development, academic achievement, early ages.

I. Introduction

Any evaluation activity has as final purpose a decision: on one side the appreciation of results expressed in terms like success versus failure, promoted versus unpromoted etc., and on the other side the improvement and optimization of the process which led to the respective results involving differentiation, individualization, change of strategy, means, methods etc.

Considering that academic evaluation represents the process by which we delimit, obtain and provide useful information about the child's developmental level and implicitly information about the acquisitions achieved, which allow subsequent decisions, we consider this issue as highly useful in the context of new educational policies – development of pupil's/student's personality- centred so that he might be capable of easy adaptation and integration in the sociocultural environment he lives.

¹ University of Pitești, Faculty of Education Sciences, smariamagdalena@yahoo.com, magdalena.stan@upit.ro.

The issue of academic evaluation from the perspective of psychical development and of academic achievements represents a challenge having multiple theoretical, methodological and practical-operating implications.

The identification of the learning outcomes in successful terms imposes both their measuring and appreciation. The intrinsic measuring represents an evaluative quantitative approach through setting the accumulations at informational and formal (competences, capacities, skills, etc.) level. The appreciation implies the reference of the obtained data to a set of criteria and norms, thus establishing the value of the outcomes. Therefore, appreciation becomes a qualitative description of outcomes. Appreciation involves data processing according to the objectives of evaluation, emitting value judgments upon outcomes and drawing conclusions for decision making. Thereby, evaluation targets a correct but approximate appreciation, having a lower degree of objectivity (M. Manolescu, 2008, p.85).

The difficulty to ensure high objectivity lies mainly in the reference to the system of value criteria which derive from the onset objectives (framework objectives). Also, we have to bring up the fact that evaluation is always relative since it refers either to a norm or to criteria. When these references change, the evaluation result modifies. (M. Manolescu, 2008, p.242). The author continues the analysis specifically, mentioning that in current practice some evaluations performed by educators fall into a precise criteria system, in a well-designed reference framework while others relate to ambiguous reference systems, poorly outlined, thus affecting the *validity* of evaluation tests and experiments.

We have to mention that the tests with a high level of objectivity (the concordance level among the appreciations of several evaluators) and validity are represented by the standardized tests.

The current practice demonstrates that the evaluation at the end of a cycle of studies and particularly, at the end of a stage of the cycle of studies, is not carried out unitarily and objectively, each primary school teacher settles his evaluative tasks with certain difficulty levels and with slight differences in complexity. The evaluation instruments elaborated by the teacher have low technical qualities especially *fidelity*. Admittedly, fidelity is related to the quality of measuring, in this case the grading form being very important.

Another poorly represented characteristic of the evaluation instruments performed by the teacher is *sensibility*- the capacity of the evaluation instrument to differentiate pupils' outcomes. In the specialty literature, it is considered that sensibility does not represent the most

serious problem of academic evaluation (M. Manolescu, 2008, p. 244), but in evaluating academic outcomes „some use only the extremities while others the middle part” (Y. Abernot, in *ibidem*) of the evaluation scale. The problem is more delicate in the qualification grading system, where pupils’ grading is mostly around the qualification *very well* and *well*, the qualificative *sufficient* or *insufficient* being very rarely allotted. As a conclusion, we can say that the academic outcomes will be significant only if measured with an evaluation instrument which keeps to the technical qualities of standardization.

II. Methodological aspects of psychological development and academic achievement evaluation

The methods and instruments by which precise, quantifiable and objective data about the subjects are measured and collected in a relatively short period, underlining the differences among them and which offer a diagnosis and a prognosis about their future evolution are the *tests*. If initially, tests had measured only children’s intellectual development, later on they had been extended to determining aptitudes for professional choice and nowadays they are used in relation to all psychological functions and field of activities.

In the psychopedagogical research as well as in the educational practice of children’s/pupils’ knowing and evaluation they are used in determining the level of psychological development and to their academic selection.

According to the field of research they belong to, tests can be (Holban, I., 1995):

- *Pedagogical tests* (of academic achievement, docimologic or performance) are destined to the study of pupils’ academic achievement;
- *Psychological tests* destined to the investigation of psychological development – cognitive, affective, psycho-motor – of the educational subjects;
- *Sociometric tests* destined to the analysis of the academic group structure and of the interpersonal relations within the group.

➤

The main attributes of a pedagogical standardized or nonstandardized test are: *validity*, *fidelity*, *objectivity* and *applicability* (Stoica, A., 2001).

Validity refers to „the fact if the test measures what it is purported to measure” (Ausubel, D., Robinson, F., 1981).

a. content validity expresses the measure in which the tests cover a representative sample of the elements to be tested. It is appreciated by a panel of „experts” (specialists in the field, teachers with didactic expertise etc.), by estimating the concordance between the test items and learning outcomes specified through the test **objectives** defined previously to item elaboration.

b. construct validity expresses the accuracy with which the test measures a certain „construct” indicator (creativity, motivation etc.).

c. concurrent validity refers to the correlation among the results obtained by a pupil to a test and „some similar behaviours” (Ausubel, D., Robinson, F., 1981). For instance, if a pupil obtained good results at a test measuring the capacity to use the four mathematical operations and does not commit arithmetical errors at a problem solving test which uses such operations, we say that the first test has a concurrent validity.

d. predictive validity refers to the degree to which the test can predict the pupil’s future outcomes. For the estimation of such type of validity, a correlation is established between the results obtained at the predictive test and another criterion which measures the same behaviours (the results of a mathematical test will be correlated to the academic outcomes obtained by the pupil at the discipline mathematics).

Fidelity represents the quality of a test to produce invariable results during its repeated application. It is influenced by the length of the test (the longer the test the higher fidelity), by score dispersion, (the larger the dispersion the higher fidelity), objectivity of the test (the test made up of objective items has a higher fidelity), by the grading system (an ambiguous grading scheme changes the test fidelity).

Objectivity of a pedagogical test refers to the concordance level among the appreciations made by the independent „experts” concerning a good answer for each test item.

Applicability represents the quality of the pedagogical test to be administered and interpreted easily. The selection criteria for tests with a high level of applicability are (Ausubel, D., Robinson, F., 1981): the importance of the content to be measured, the concordance between the test content and form and the pupil’s age, the cost and time necessary for administration, objectivity in grading and interpreting the results.

II.1. The nature of academic achievement tests

Academic achievement tests are devised to measure the effects of educational programs. Admittedly, there is a difference between the achievement tests and the aptitude tests, the latter including the general intelligence tests, the special aptitude tests and the multiple aptitudes batteries. Anastasi A. (1969, p. 391) emphasizes this differentiation, by mentioning:

✓ The difference between the achievement tests and the aptitude tests is the difference in the uniformity level of the relevant previous experience, mentioning that the achievement tests measure the relatively standardized effects of the accumulated sets of experiences through specific instructional programs: the elementary English course, reading, mathematics, while the performances in the aptitude tests reflect the cumulative influence of multiple experiences of the everyday life. Thus, the aptitude tests measure the learning effects in unknown and uncontrolled conditions while achievement tests measure learning effects which took place in partially known and controlled conditions.

✓ The differentiation between the two categories of tests can be made from the point of view of their use, therefore the aptitude tests being used as predictors of subsequent performances. Achievement tests represent in general, an evaluation of the individual status at the end of an educational program. The emphasis is laid here on what the individual can do at the respective moment. His difference is probably more precisely illustrated by the validity estimation procedures of achievement tests in contrast with those of the aptitude tests validity. Although the validity of a criterium represents the most direct way in establishing the aptitude tests, the achievement tests are evaluated in terms of content validity.

However, there cannot be a rigid differentiation between the aptitude tests and the achievement tests. Some aptitude tests may depend to a great extent on initial specific and uniform learning, while some achievement tests cover certain general and nonstandardized educational experiences relatively. Thus, the achievement tests can be used as a predictor of future academic learning. For instance, a pupil's progress in arithmetics determined by scores at an achievement test can be a predictor for a future success in field of algebra. Thus, achievement tests can offer efficient indices of a future performance.

A concern the differentiation between the achievement tests and the aptitude tests we should be cautious in the opinion that achievement

tests measure the learning effects and the aptitude tests measure the inborn capacity, irrespective of learning. All psychological tests measure an individual current behaviour which reflects inevitably the influence of previous learning. The tests which measure the outcomes of the previous learning can be under certain conditions predictors of the future leaning.

A central concept of psychometry which covers both the category of aptitudes and the category of achievements is the ***development of abilities***. All tests of abilities (no matter whether it represent general intelligence, special aptitudes or even achievements) measure the developmental level attained by the individual at one or more abilities, without being able to specify how or why he reached the respective level.

Anastasi A. (1969, p. 392) includes in the category of achievement tests, the tests oriented on the understanding and application of scientific concepts, literary interpretation or art appreciation; also this category includes tests of intellectual aptitudes which influence individual performance in a wide range of activities such as study aptitudes: interpretation of tables, the use of dictionaries and indexes, text deciphering, arithmetical operation.

At a general level, achievement tests can be projected to measure the effects of education at the level of logical thinking, of critical evaluation of conclusions, of the problem solving techniques.

Mitrofan N. (2010, p.29) defines ***achievement tests*** as instruments „which monitor the individual diagnosis of the achievement obtained by those involved in the educational process, or the acquired information, skills and capacities. It is the informational aspect (achievement of knowledge), and the formative one (capacity to operate, to act)”.

The achievement tests or the knowledge tests (in Albu. M., 1998, p. 72) represents a way of appreciating the stock of information of a person. After Thorndike, Hagen, 1961 (in Albu. M., 1998, p. 72) the main difference between the two categories of tests is given by the direction of the examiner's interests. He applies a test of aptitudes if he wants to know what a person can learn and use a test of knowledge if he wants to know what the person has learned until the moment of examination.

II.2. The use and functions of academic achievement tests

The main characteristic of the achievement tests is represented by the degree of objectivity and uniformity. If they are built correctly, they have the advantage of the adequacy of the covered content and the

decrease of irrelevant operations and factors in marking the procedures (Anastasi A., 1969, p. 392).

The tests of academic achievement present an important property – the retrieval of educational programs. These are used at identifying children with special educational needs as well as in measuring their progress. Another characteristic of these instruments refers to the effect they have in easing learning by their periodical administration. These tests indicate the deficiencies; offer the direction of the future learning and motivational support for the learner. The value of the „knowledge result” stimulus had been demonstrated repeatedly by psychological experiments in different learning situations with subjects of different ages and educational levels. When the examinations of achievement are used as a learning process facilitator, it is desirable that pupils should know their deficiencies immediately after the test.

The achievement tests offer means of adapting the individual instruction needs. Instruction can be fruitful when it satisfies the learner’s needs regardless of its level. In the efficiency of teaching it is necessary first to establish the individual level of what can he do and what does he know about the respective issue. By applying tests at the beginning of each academic year important steps can be taken to fill in major deficiencies from the system of knowledge, tracked down by these means.

The achievement tests can be used as a support in the evaluation and improvement of the instructional programs, as well as in the elaboration of educational objectives. These can offer information about the adequacy of the essential content and of the targeted skills. Also, they can indicate the percentage of the acquired course content and its persistence in time.

A tests of achievements rigorously built is an efficient way to illustrate the concrete behavioural changes which can be made through an educational experience.

II.3. Classification of academic achievement tests:

According to their **goal**, the academic achievement tests can be:

- **Standardized achievement tests:** elaborated by specialists, administered, quoted and interpreted under standardized conditions, destined to be widely used; these are used in order to compare the activity of a group with another group activity, to compare the results of a class or of an individual to external data offered by the test norms.

- **Informative achievement tests:** elaborated by the teacher of the class and interpreted in relation to the scores of the respective class; these are mainly used to measure the attainment of the immediate class goals and are administered to the group for which they had been elaborated.

The academic achievement tests can perform different *functions*, and can be classified as it follows:

- **Absolute criterion achievement tests** (*mastery tests*): verifies if subjects have absolute necessary knowledge for a certain goal; these evaluate the individual performance as related to a previously specified performance level; the minimal performance is reported and the subjects must complete at least the absolute criterion fixed stipulations.

- **Relative criterion achievement tests** (*survey tests*): help to build a hierarchy of subjects according to their knowledge; relate to maximal performance, the subjects being hierarchized according to the score obtained at the test.

- **Diagnosis achievement tests** (*diagnostic tests*): their role is to identify the subjects' deficiencies, absent or incomplete knowledge; they aim at tracking down the domains where performance is very low.

II.4. The elaboration of academic achievement tests

Roughly speaking, the elaboration of a test or of a psychological experiment develops over several stages, non-sequentially, with repeated recursions until the attainment of desired qualities (fidelity, validity, power of discrimination, difficulty). M. Albu (1998, p. 199) mentions that test items are formulated according to the goal of the elaboration (theoretical or practical). The group of items is administered to a sample of persons extracted from the population destined for the test; answers obtained at every item are analyzed, inadequate items are identified and excluded. The next stage is the error causes analysis (related to items, sample of subjects or to the way items are administered). Faultless items are introduced in the second part of the item analysis which applies procedures specific for the chosen method.

Test characteristics are calculated: fidelity, validity, power of discrimination and if these are acceptable, the test is administered to a sample of person similar to the first one, in order to verify if the mentioned qualities are preserved.

If these hold, we can determine the test norms and establish the interpretation rules of the scores.

1. *Test defining* refers to „the multitude of behaviours which can be used to measure the specific attribute or the characteristic the test refers to”. (Murphy, Davidshofer, in M. Minulescu, 2003, p. 79). The behaviours supposed to be expressions of subjects' performance or knowledge are maintained. In case of achievement tests, a table of specification is made which indicates the item distribution according to the test behavioural objectives and to the themes to be verified. The table of specification serves as a base for the nature of the items included in the test, as each cell of the table matches certain types of items.

2. *The types of items* are chosen in accordance with the stipulations of didactic objectives, with the learning contents so that their nature might allow the appreciation of the information processing capacities and of applying the intellectual abilities formed under different circumstances. The sequence of items in the tests can be aleatory or can be done according to their difficulty, in an ascending or descending order. The items will be thus placed so that the content of an item should not suggest the answer of another set in its proximity.

3. In order to ensure a better value to the achievement tests N. Mitrofan (2005; 2010, p. 102) quoting P. Kline (1993) considers that an important stage is the modality of elaborating items: the obligation to cover the integral content of the training courses; the number of items must be considerable in order to ensure fidelity and scoring must be very objective; the tests should be valid and contain informational and operational aspects. *Creation of the item bank*: items which are introduced in the data bank can come from several sources: are formulated by psychologists based on a theory (if the test measures a psychological construct); are elaborated by specialists of the field where the test will be applied; are elaborated by diverse persons as illustrations of the behaviours they consider characteristic for a certain construct; are extracted from existing tests, which had been administered to a sample similar to the sample destined for the new test. The next stage is the determination of the quotation manner.

4. *Construction of a test variant*

In order to elaborate a test variant it is necessary to set its length. The length of a test is chosen according to the unallotted time for test administration, which is settled in general taking into account the conditions of its usage. This depends on: age of subjects, subjects' reading skill, the length and difficulty of items, the wording.

5. *Test administration*: is elaborated under the same conditions in which it will be used in practice. Generally speaking, if the test is

administered in the class, the pupils should be informed in time, when and where the will the test take place and its subject matter.

6. *Analysis of the test results*: we verify if the administration instructions had been complete and if the working time was sufficient. The answers are then processed through the following: item analysis, disposal of inadequate items, fidelity, validity, etc.

7. *Test formulation proper*.

One of the most used procedures in building a pedagogical test is the specification matrix.

Specification matrix is a double entrance table in which content areas to be measured are shown on the horizontal and cognitive levels according to which we measure these contents vertically (for instance: according to Bloom's taxonomy of educational objectives we can insert columns which refer to knowledge, understanding, application, analysis, synthesis vertically). The table cells placed at the intersection of content elements and taxonomic levels for the cognitive level include the item percentage (from the total items of the test) which will be used in test elaboration.

In elaborating the specification matrix, the following steps are made: we identify the types of behaviours/fields which will be tested (knowledge and understanding, application, analysis, synthesis, problem solving); we identify the content elements which will be tested; we determine the balance on each domain and content element, the each domain will have within the test; we complete the matrix cells by multiplying the data on the lines with the respective columns.

Conclusions

In evaluating the performance level attained by pupils in their academic activity and in order to realize a comparison among pupils' results, the use of standardized (tests) instruments proves to be highly necessary. Definitely, the use of a standardized instrument in academic evaluation does not solve all problems related to this activity, but we can be sure that it offers a higher degree of objectivity and unity with a view to know, appreciate, and interpret the learning outcomes, to control the future educational actions and to prepare educational decisions.

Bibliographical references

- Albu, M. (1998). *Construirea și utilizarea testelor psihologice*. Cluj Napoca: Editura Clusium.
- Anastasi, A. (1969). *Psychological Testing*. NY: MacMillan Company.
- Crețu, T. (2001). *Psihologia vârstelor*. București: Editura Credis.
- Cucoș, C. (2006). *Pedagogie*. Iași: Editura Polirom.
- DiPerna, J.C., Volpe, R.J., & Elliott, S.N. (2002). A model of academic enablers and elementary reading/language arts achievement. *School Psychology Review*, 31, 298-312.
- Duncan, G., et al. (2007). School Readiness and Later Achievement. *Developmental Psychology*, 43(6), 1428-1446. <http://dx.doi.org/10.1037/0012-1649.43.6.1428>
- Golu, P., Golu, I. (2003). *Psihologie educațională*. București: Editura Miron.
- Ionescu, T., Benga, O. (2007). Reconceptualizing early education scientific grounds: school readiness in focus. *Cognition, brain, behaviour*, XI(1), pp. 49-65.
- Manolescu, M. (2008). *Evaluarea școlară. Metode, tehnici, instrumente*. București: Meteor Press.
- Manolescu, M. (2006). *Curriculum pentru învățământul primar și preșcolar*. București: Editura Credis.
- Meisels, S.J. (1999). *Assessing Readiness*, Report CIERA
- Mitrofan, N. (2009). *Testarea psihologică. Aspecte teoretice și practice*. Iași: Editura Polirom.
- Miclea, M., Mihalca L. (2007). A computerized platform for the assessment of school readiness. *Cognition, brain, behaviour*, XI(1), 83-90.
- Mih, V. (2010). *Psihologie educațională*. Cluj Napoca: ASCR.
- Minulescu, M. (2003). *Teorie și practică în psihodiagnoză*. București: Editura Fundației România de Mâine.
- Muntean, A. (2006). *Psihologia dezvoltării umane*. Iași: Editura Polirom.
- Papalia, D. E., Wendkos Olds, S., Duskin Feldman, R. (2010). *Dezvoltarea umană*. București: Editura Trei.
- Păiși Lăzărescu, M. (2008). *Psihologia vârstelor*. București: Editura V&I Integral.
- Radu, I. T. (2008). *Evaluarea în procesul didactic*. București: EDP.
- Sillamy, N. (2009). *Dicționar de psihologie*. București: Editura Univers Enciclopedic.
- Sassu, R. (2007). The evaluation of school readiness for 5-8 years old children- cognitive, social –emotional and motor and physical health perspectives. *Cognition, brain, behaviour*, XI(1), 67-81.

Stoica, A. (2003). *Evaluarea progresului școlar. De la teorie la practică*. București: Humanitas Educațional.

*** National Education Goals Panel (1997). *Getting a good start in school*. <http://govinfo.library.unt.edu>, accesat la 04.09.2008