Dynamic Assessment of Students’ Mathematical Competencies for Lower Secondary Education (Grades 5 and 6)

Carolina Staiger
University of Education Weingarten,
Kirchplatz 2, D-88250 Weingarten
staiger@ph-weingarten.de

Abstract
Dynamic assessment constitutes an effective alternative to the conventional test which records only the current level of performance. Thereby researchers want to ascertain not only the current level of competence of a person, but particularly a person’s learning ability which is evoked by feedback (e.g. elaborated feedback). Thus they can achieve a more comprehensive assessment of common performance. The current state of our work focuses on the one hand the theoretical background of dynamic assessment and on the other hand the ambition to develop approaches concerning the development of a dynamic testing method in the domain of mathematics in lower secondary education in grades 5 and 6.

Introduction
The idea of dynamic assessment has a long tradition. In 1964 Wygotski was already convinced that for assessing a person’s performance it is necessary to record the learning ability in addition to the current level of development (cf. Wygotski, 1964). Nevertheless it seems that researchers in the domain of performance diagnostics still prefer status tests. Furthermore this alternative testing method especially in the domain of mathematics is largely immature. We want to counteract this deficit with the development and testing of a dynamic testing method to assess mathematical competencies of students in lower secondary education in grades 5 and 6.
The article should be a first step in this direction. Thereby it can be divided into two main parts. In the first part, the theoretical background will be clarified.
In the second part first approaches in the field of mathematics will be worked out.

Dynamic Assessment
Dynamic assessment has its roots in intelligence diagnostics. The central idea goes back to the Russian psychologist Lew Wygotski. According to him in order to assess a person’s performance it is necessary to ascertain the current level of development and in addition the learning ability (“zone of proximal development”), thus the ability of an individual to reach improvement under supportive conditions. (cf. Wygotski, 1964, S. 212)
In other words it is of interest to ascertain not only the current level of competence of a person, but particularly the person’s learning ability which is evoked by feedback (e.g. elaborated feedback). Dynamic assessment can be classified into the domain of process diagnostics. Thereby the modification of a variable (here: performance) can be recorded, because of interim interventions at two or more different dates. In this connection it is not only interesting to get a before and after comparison, but particularly to get information about the learning process. Compared with status tests, dynamic assessment intends to record not only the current level of performance, but especially a person’s learning ability. With regard to a status test a person’s actual state shall be identified. More precise competencies that someone had learned in the past are recorded. (cf. Pawlik, 1982)
Let us suppose that we test two different students; one of them reaches a good result, the other one a bad result. Does that mean that the weaker student has less understanding in this domain than the person with the good result? To answer this question we should analyse his or her learning behaviour in the past few years. Did he work diligently and did he learn intensive, but nevertheless...
he could not reach outstanding results? Or was he lackadaisically and neglected school? In the second case he or she would perceive to be too low in spite of its possibly high learning ability. In this way it is impossible to get information about a person’s learning ability, but only about the current level of competencies. On the contrary process diagnostics intends to record modifications of a variable (e.g. performance). Consequently in dynamic assessment the performance within a test procedure is ascertained which includes the current level of competence and the learning ability. Thereby each person is subject to a learning process, whereby his or her competencies can be enlarged and thus can meet all demands easier. The extent of gaining competencies shows the degree of learning ability.

Performance
Performance is the directly observable behaviour when working on a definable class of requirements (e.g. multiplication, arithmetic, curricular guidelines for lesson) (cf. Hasselhorn; Marx & Schneider, 2005). The current level of performance is the present cognitive skills which a person could acquire from past learning processes. The terms “learning ability” or “learning potential” mean the “zone of proximal development”. Furthermore it is the extent of expandability of cognitive skills of a person. (cf. Dörfler et al., in press, p. 3)

How can I record the current level of performance and the learning ability at the same time? To ascertain performance you need an indicator. In the field of performance studies you use items. In the domain of mathematics these are mathematical calculations, whereby each item has a different level of difficulty. So they can be classified in a certain level within a scale. Therefore in the case of recording the current level of performance it is important, how many items a person could solve and which level of difficulty they solve it at. Thus it is possible to attach a person to a certain level of competence.

According to the assessment of the learning ability the following questions are important: How many prompts did a person need to solve an item? To what extent has a person gained from the interventions?

The use of recording learning ability
Which advantages does dynamic assessment have, respectively which advantages does the assessment of the learning ability have? Concerning the value Embretson mentions three significant aspects: first it is possible to ascertain skills better, second new and inaccessible skills can be identified and third skills can be fostered. (cf. Klauer, 1994)

Test format
Within the dynamic assessment there are two different test formats, the “test-train-test” (“long-term learning test”) and the “train-within-test” (“short-term learning test”) (Guthke & Wiedl, 1996, p. 92).

With regard to our intention to assess the learning ability of students, we chose the train-within-test format. Thereby the test procedure and the instructions are synchronized. That involves a temporal advantage. If a person gives the wrong answer to an item, thus an adequate feedback will follow that includes a strategy which can help solving the item correctly. In the course of this the intervention procedure itself is estimated, e.g. how much support does a certain person need to solve the item correctly? Researchers observe how a person can use these provided strategies. In contrast to the test-train-test, whereby the focus is on fostering participants, the train-within-test focuses on efficient diagnostics of the learning ability, so at this point the aspect of measuring is central. So the train-within-test format makes it possible to ascertain a person’s direct reaction to provided feedback. (cf. Dörfler et al., in press)
But in this context certain challenges arise, such as the measurement of performance modification and the construction of short, concise and particularly (concerning the items) suitable feedback, because there is no time for extensive interventions.

**Interventions**

**Feedback**

Dynamic assessment includes instructions in the form of feedback and (meta-) cognitive strategies. After answering an item the computer software gives feedback. If the answer is wrong the participant is provided suitable strategies and/or solution-oriented prompts. Feedback gives information about one’s performance in a test situation, intended to affect regulating on prospective process.

We chose the following types of feedback:

Knowledge Of Result (KOR) informs the student, whether his or her answer is right or wrong. In case of a wrong answer, the student doesn’t get the right one.

Knowledge Of Correct Result (KCR) informs the student of the right answer after answering the item.

Elaborated Feedback comprised of KOR and KCR, gives error-specific hints for corrections or references to cognitive and meta-cognitive respectively strategies and should have the function to help students to solve an item. (cf. Narciss, 2006)

*(Meta-) cognitive strategies*

Learning strategies are intentionally and purposefully behaviour that help to master items. Konrad distinguish three levels of strategies, “cognitive learning strategies”, “meta-cognitive strategies” and “resource-based strategies” (Konrad, 2008, pp. 50-51).

In our project only the first two levels are interesting for us. Additionally we want to pick up heuristics and its repertoire of strategies and tools. In dynamic assessment strategies should serve as tools to solve the given items easier.

**Feedback structure**

We developed an interim model of feedback structure which shows how the given types of feedback can be realized.
Altogether, a student can get up to four responses. After each wrong answer he or she will receive one particular type of feedback. In the case of the first wrong answer the student will be informed that his or her answer is wrong (level 1), so he or she can rethink his or her own approach. If the student can’t solve it afterwards, he or she will be provided a strategy or solution-oriented prompt (level 2a/b). Thereby the Elaborated Feedback on level 2b – which will be given after three wrong answers – will be more concrete than the previous one on level 2a. For example: The student will receive the first step to solve the item. Finally, if he or she gives the fourth wrong reply, the correct answer will take place (level 3).

Research objectives and questions

Research objectives

We intend to develop a computer-based dynamic testing method to record mathematical competencies of students in grades 5 and 6 of secondary level 1. Furthermore we want to design items whereby we can assess the particular level of competence and learning ability of a student in different mathematical sub-disciplines. In this connection we will revert to results of preliminary work to the topic acquisition of mathematical development of performance in secondary education (cf. Kleine, 2004). Besides we will develop suitable feedback to each item which helps to solve an item easier.

Research questions

Questions we want to clarify are:
To what extent can you assess the learning ability of students in mathematics?
How can you develop test instruments which can assess the learning ability by means of responses?
Which types of items can be advantageous to assess the learning ability?
Which types of feedback and strategies are most effective to support inferential processes, when solving a mathematical item?
How can you implement the test method to get objective, reliable and valid results?

The development of the test

Working Plan

Concerning the development of the testing method we developed a working plan which shall give a rough overview of our project.
This year we design specific items to the different sub-disciplines. These items must map the learning process, cover low and high hierarchical range of requirements, must offer the possibility to implement feedback and strategies and furthermore they should be rather difficult than easy to make certain interventions possible. In addition we will choose adequate types of feedback and (meta-) cognitive strategies which can be applied in connection with the given items. Afterwards we will supply these items and feedback into a specific computer program.

As a second step, we will arrange a field study with a small test group. Therefore we will choose 20 to 30 suitable students in grades 5 and 6 which were tested in advance for previous knowledge, motivation and intelligence. In a pre-test we will do cognitive interviews with the chosen students. Thereby participants should give feedback in form of thinking aloud concerning the adequacy of items.

As a third step, we will organize our main study, whereby we want to choose about 200 students in grades 5 and 6 from secondary level 1. For test instruments we will apply our developed computer-based items with integrated Feedback. Finally, we will analyse and publish our results.

**The Scaling of the performance data**

*The dichotomous Rasch-Model*

The items, which will be developed, should map a certain level of performance respectively the level of competence. Items of varying difficulty are positioned on a continuous scale of competence which is divided in different levels of competence (cf. Schweizer, 2006). By means of modelling the response probability, you can estimate the item-difficulty and later you can assess a person’s level of competence.

The Rasch-Model admits an assumption of connections between latent variables and the probability for the appearance of certain answers. The significant variables person’s ability ($\theta$) and item-difficulty ($\delta$) are independent from each other, but both depend on the answer of a test person. The behaviour of a participant on an item is a categorical variable with the value 0 (wrong) and 1 (right). The intention is to get an indicator about the characteristics of level of competence to be measured due to the number of solved items. (cf. Kleine, 2004) With incremental capability value the probability of solving should rise continuous and monotonous.

![Fig. 4: Item characteristic curves](image-url)
**Outlook**

On the basis of the compiled theoretical background and approaches to implementing, furthermore we will develop low and high hierarchical items and suitable feedback. Thereby we will use scientific results about mistakes students make in certain domains. At first we want to focus on algebraic subdomains (e.g. fractions) and on this we will differentiate the created model of feedback structure. In this connection this model should serve as a rough guideline and should not be seen as fixed. Finally, in the development of the item-feedback-composition it is important to keep in mind that this must be realisable with a specific computer program.

Concerning the computer-based test method there are also some objections which should be clarified:
Which factors can influence a computer-based dynamic assessment and distort the results?
How can we counteract these interfering factors?
And is the computer program able to respond individually to a student’s mistakes and problems?
And if so, how does it work?
The documentation of the pre-test (step 2) and finally the adherence and analysis of the planned random sample (step 3 and 4), will become part of my dissertation.
References