Dynamic Assessment and its Potential for the Assessment of Reading Competence [1, 2]

Tobias Dörfler, Stefanie Golke, and Cordula Artelt

Abstract This article describes the approach of dynamic assessment, focusing on general approaches as well as specific constraints for the assessment of reading competence. Starting with an overview of the literature on dynamic assessments within educational research, the framework of dynamic assessment in which the current level of competence and (domain-specific) learning ability are assessed, is discussed with regards to its methodological and diagnostic implications. Reading competence is introduced as a domain in which interventions prove successful, and as a domain principally suitable for the assessment of learning ability. Furthermore, it is discussed whether elaborated feedback given within the procedure of reading competence assessment does uncover a learning ability which is specific to the domain of reading competence or not.Times NR 10

Keywords *dynamic assessment, learning ability, reading competence, text comprehension, feedback, IRT models*

Dynamic Assessments in Educational Research

Educational research aims to investigate competencies and their development and modifiability. To this end, scientists and practitioners are in need of theoretically and empirically well-established competence models as a basis for appropriate operationalizations and the development of assessment techniques. Dynamic assessments are one approach to gaining insight into the current level of competence as well as into how this competence can be influenced by specific educational interventions. Advocates of the dynamic assessment approach argue that these assessments are able to provide "more reliable, valid, and diagnostically and prescriptively useful estimates of the tested abilities, or competence, than would be the case if the same tests were administered under traditional psychometric methods." (Dillon, 1997, p. 164).

Different Approaches of Dynamic Assessment

According to Lussier and Swanson (2005, p. 66) dynamic assessment is a "procedure that attempts to modify performance, via examiners assistance, in an effort to understand learning potential". Traditional diagnostic strategies in the broader framework of dynamic assessment are the assessment of structural cognitive modifiability (see Feuerstein, Rand, Haywood, Hoffmann, & Jensen, 1983), learning potential assessment (see Budoff, 1987a), learning test (see Guthke, 1982) or testing-the-limits (see Carlson & Wiedl, 1979). As it is beyond the scope of the article to describe in detail these assessment strategies, the reader is referred to the comprehensive reviews of Guthke and Wiedl (1996) or Poehner (2008).

While dynamic assessments differ in the names they carry, they share a common purpose in approaching the assessment of learning potential. Supporters of dynamic assessment adhere to the concept by Vygotsky (1964) who stressed the importance of the "zone of present development" and the "zone of proximal development". The zone of present development represents the performance level an individual can achieve without external assistance. As a sideline, this unassisted performance can just as well be assessed by means of traditional (static) tests. Furthermore, Vygotsky has pointed out that the individual might improve its performance under guidance provided by adults, more capable peers or a computer. The distance between current developmental level achieved without assistance and the level of potential development ascertained through guidance is defined as zone of proximal development. It provides additional information about individuals' learning abilities which are assumed to manifest in upcoming developmental steps and indicates their capacity for cognitive modifiability.

Dynamic assessment integrates the concept of the zone of proximal development, i.e. the approach to examine both the performance without assistance and additionally achievement under guidance. Within dynamic tests, often a computer-based

1) This research was supported by grant AR 307/7-1 from the German Research Foundation (DFG
2) Accepted Author's manuscript. This article was published in:
Studies in Educational Evaluation, 35 (2009) 77-82, doi: 10.1016/j.stueduc.2009.10.005

(tutorial) system takes the part of providing the individual with help drawing on diverse interventions by means of aids, instructions, feedbacks, or prompts. It is apparent that the usage of assistance within dynamic assessments is its unique characteristic. The crucial point within dynamic assessments is to design the best possible guidance within a certain dynamic test in order to exploit individuals' full potential. Therefore, the construction of help needs to consider at least the competence domain, test format, and students' characteristics, as will be outlined below.

Formats of Dynamic Assessment

Two formats of dynamic assessment are commonly applied (see Dillon, 1997), both of which use educational interventions (e.g., instructions, feedback) to induce performance improvement. The extent of provided help differs by the formats of dynamic assessment. The first, most frequently employed format is a *test-train-test design* with training occurring between a pre- and a post-test. Following administration of pre-test items, participants are taught the most useful strategies for item solving. In a post-test, parallel items determine the extent to which participants have improved their performance as a result of training and their individual learning abilities. This test design allows on the one hand for extensive interventions given that separate training sessions are established. On the other hand, the whole assessment procedure is rather time-consuming since pre- and post-tests are normally administered on different days. The second format comprises a *train-within-test design* and the following test procedure: If items are solved incorrectly, a particular kind of intervention (e.g., feedback, aid) is provided immediately after failure. Due to this connectivity of intervention and assessment, interventions within train-within-tests designs are rather limited in time. As a consequence, the interventions need to be terser and less complex than those within test-train-test designs, but should still maintain specific to the task demands of the test at hand. Therefore instructional hints, prompts, or feedback are applied that are suitable for the focused domain. Figure 1 provides a schematic overview of the test-train-test and the train-within-test designs.

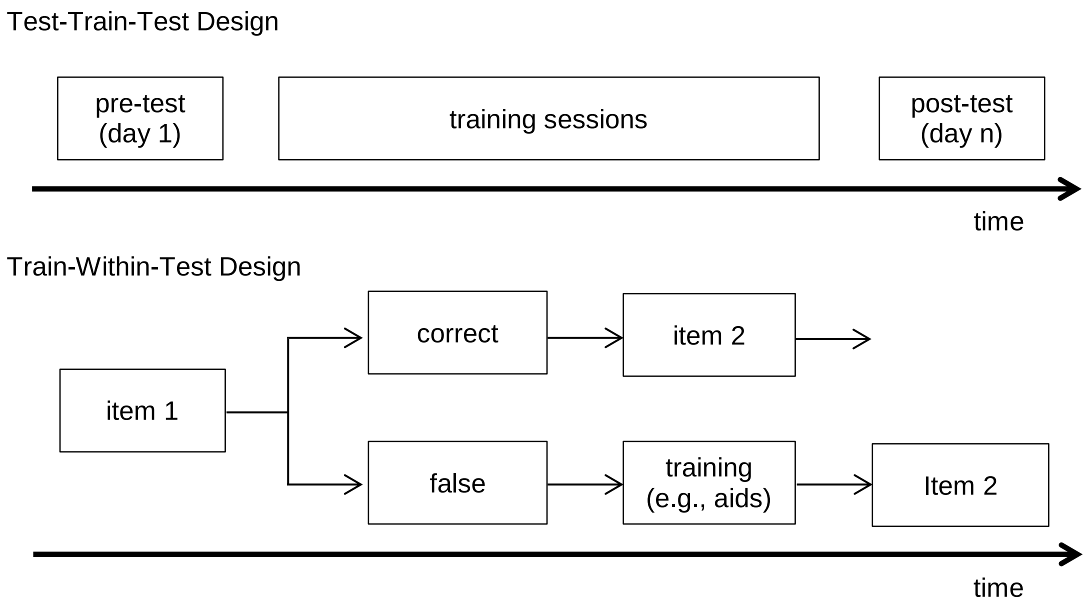

Figure 1. Traditional formats of dynamic assessments. Test-train-test format (above) with a pre-post-test design and separate training sessions and train-within-test format (below) that only has one session with training parts included.

Validity of dynamic assessments

There is evidence that dynamic assessment approaches bear potential for application in everyday diagnostic practice (Elliott, 2000). The following results characterize major findings on validity issues of dynamic assessments, as also reported in meta-analytic reviews (see Lussier & Swanson, 2005). Studies addressing the validity of dynamic tests reveal a small but mentionable superiority of a number of different criteria (e.g., school performance) over static tests (cf. J. F. Beckmann, 2001; Budoff, 1987b; Carlson & Wiedl, 2000; Guthke & Wiedl, 1996). Concurrent and prognostic validity in the field of intelligence testing has been found to be slightly higher for post-tests after intervention in a dynamic test-train-test design (Flammer, 1975) and incremental validity can be gained using a second measure (Guthke & Stein, 1996). Carlson and Wiedl (1980) demonstrated that while static intelligence tests are a fairly good predictor of school achievement, dynamic intelligence tests are even better.

This finding is also supported by more recent work in the science domain (Hessels, in press).

In contrast to status tests, dynamic assessments are commonly affirmed to be less biased against ethnic or social minorities (Hessels, 2000).

Given the research reported so far, dynamic assessments provide advantages for the terms of the validity of test results. But they also bear challenges. This particularly applies to issues of scoring and scaling procedures which are problematic due to individualization. The following section focuses more closely on the methodological challenges of dynamic assessment, especially for train-within-tests.

Methodological Challenges of Dynamic Assessment

The crucial issue regarding scoring and scaling of response data from dynamic assessment concerns how to score performance that reflects changes induced by the dynamic assessment procedure itself. While test-train-test approaches assume that change occurs after an intervention phase, researchers engaged in train-within-test assessments are often required to score the intervention procedure itself (Embretson, 1987; Guthke & Wiedl, 1996), such as recording the correctness of responses and the number of provided aids that were needed to reach a specific criterion (J. F. Beckmann, 2001; J. F. Beckmann & Guthke, 1999; Campione & Brown, 1987). The fewer the number of aids, the higher the person's learning ability, because s/he requires only minimal assistance in solving items. Thus, the learning ability has to be assessed in addition to the initial ability level (Klauer, Kauf, & Sydow, 1994; Klauer & Sydow, 1992).

Convenient models modeling response data from dynamic assessments are based on item response theory (IRT). These models are able to deal with many of the measurement problems associated with classical test theory, in particular in the case of dynamic assessments (Sijtsma, 1993a, 1993b). Given the basic assumption of dynamic assessment that performance is related to both initial ability level and learning ability, a two-dimensional item response model is an appropriate representation of the data from test-train-test designs. Klauer and Sydow (1992) or J.F. Beckmann and Guthke (1999) could show for train-within-test designs that a two-dimensional model with initial ability level and learning ability as correlated competencies fits the data better than a uni-dimensional model does.

In principle, the models apply the same logic: According to the Rasch model (cf. Rost, 2004, p. 119), the probability of successfully solving item i for person v can be described as a function of the person's ability θ_v and item difficulty σ_i, resulting in the following equation:

$$P_A(X_{vi} = 1|\theta_v, \sigma_i) \frac{\exp(\theta_v - \sigma_i)}{1 + \exp(\theta_v - \sigma_i)} \qquad (1)$$

In a train-within-test design, each item i within the dynamic test provides each participant with the opportunity to learn, which might lead to a change in ability score θ (Klauer & Sydow, 1992). The change (increase) is termed learning ability and denoted δ. It is moderated by a person's initial ability level, item difficulty, or provided feedback. Item- or feedback-related weights x_i and y_i are therefore included for both a person's ability as well as his or her learning ability, since each item and each provided feedback enhances performance to a highly specific degree. The probability of a spontaneous correct response can thus be defined as follows:

$$P_B(X_{vi} = 1|\theta_v, \delta_v, \sigma_i) = \frac{\exp(x_i\theta_v + y_i\delta_v - \sigma_i)}{1 + \exp(x_i\theta_v + y_i\delta_v - \sigma_i)} \qquad (2)$$

As model constraints Klauer and colleagues assume that $x_{vi} = 1$ for all items and $y_i = i - 1$ are cumulated learning opportunities. Learning ability within this model is assumed to be 1) constant for each item but to vary between subjects and 2) independent of solving items correctly without or with assistance. The superiority of model P_B over model P_A is evident (J. F. Beckmann & Guthke, 1999; Klauer, et al., 1994). If individual ability levels θ_v and learning abilities δ_v are correlated, this has an additional diagnostic impact, allowing a Fan Effect or Matthew Effect (Cook & Campbell, 1976, pp. 184-185) to be captured: When taking a dynamic test, proficient children learn most and non-proficient children learn least. Increasing initial ability differences between children can thus be modeled when applying dynamic assessments, which in turn increases test reliability as compared to a status test (Klauer, et al., 1994).

Scaling models for dynamic tests in complex, multidimensional performance domains have to account for the possible existence of multiple, domain specific learning abilities. We refer to Embretson (2000), who suggested a model which is a special form of a structured latent trait model. The multi-dimensional Rasch model for learning and change allows an initial ability vector $\boldsymbol{\theta}_v$ and one (or more) additional abilities (e.g., learning ability $\boldsymbol{\delta}_v$) to be simultaneously estimated based on a person's response data over time. The model is appropriate for learning-potential assessment. Item difficulty, which is held constant for each condition k (e.g., different types and complexities of aids), is denoted σ_i. Taking these constraints into account, the model may be defined as follows:

$$P_C\left(X_{i(k)v} = 1|\boldsymbol{\theta}_v, \sigma_i, \boldsymbol{\Lambda}_{i(k)m}\right) = \frac{\exp\left(\sum_m \lambda_{i(k)m}\theta_{vm} - \sigma_i\right)}{1 + \exp\left(\sum_m \lambda_{i(k)m}\theta_{vm} - \sigma_i\right)} \qquad (3)$$

Item and person parameters are collected into a design matrix, if items are clustered according to different dynamic assessment conditions k estimating the m abilities. The ability level of person v on ability m is weighted by $\boldsymbol{\Lambda}_{i(k)m}$ in item i under test condition k (e.g., different types and complexities of feedback). This weight can be understood as an item discrimination parameter or as a factor loading of item *i* on ability dimension *m*.

The recommended models P_B and P_C are comparable and can both be applied to "dynamic" test data. The more general model P_C suggested by Embretson can be used to represent change induced by more than one ability and is therefore advantageous if more than one abilities are examined.

Apart from the methodological issues described above, an important question with respect to the additional competence dimension, labeled "learning ability" or "learning potential" is the question of whether this ability is domain specific or not. In order to analyse potential specificity of the learning ability assessed, application of dynamic assessments beyond intelligence testing, which is the most common usage of dynamic tests, is essential and according to Guthke, Beckmann and Wiedl (2003, p. 226) feasible. For example, Berman and Graham (2002) as well as Desoete, Roeyers, Buysse, and De Clercq (2002) conducted studies in which dynamic assessments were applied to examine mathematical performance in children. Other researchers have focused upon guided learning in foreign language acquisition (Kozulin & Garb, 2002; Poehner, 2008; Schneider & Ganshow, 2000). A cognitive domain in which dynamic assessments in a train-within-test format has thus far only a few applications is reading competence; a key qualification in the management of daily routines. For the remainder of this article we will concentrate on potentials and challenges of modeling reading competence within the framework of dynamic assessment.

Dynamic Assessment of Reading Competence

Reading competence can be fostered with the help of training programs. Applications of dynamic assessment in this domain traditionally focus on training sessions which are embedded between a pre- and a post-test. Assessing reading competence in a train-within-test would provide insights into an individual's reading competence level and learning ability in this domain in one session due to a person's responsiveness to hints and feedbacks (see Haywood & Lidz, 2007). The following paragraph argues for the development of such a dynamic reading competence test with a clear focus on efficient diagnostics rather than on training competencies.

Most psychological theories assume that reading competence comprises a combination of text-based and knowledge-driven comprehension processes (cf. Kintsch, 1998).

Studies on dynamic assessments of reading competence have been strongly influenced by Campione, Brown, and colleagues who have developed a number of measures related to learning and transfer processes (Campione & Brown, 1987, 1990; Palincsar, Brown, & Campione, 1991). Palincsar and Brown (1984) developed a reciprocal teaching reading intervention program in which teachers (and later students in the role of teachers) provide feedback and meta-cognitive prompts that help students' apply text-processing strategies that enhance their competence development. Later, Campione and Brown (1987) were among the pioneers who applied the framework of dynamic assessments to specific educational domains. The principle of their assessment model is to investigate how students learn from one another and how flexible they are in applying what they have learned. Results indicate that predictions based on initial performance significantly underestimated what children could achieve with minimal assistance. Thus, dynamic measures often provide better estimates of reading competence than static tests.

Prototypes of Test-Train-Tests of Reading Competence

One of the few studies to have applied dynamic assessments to the reading domain was conducted by Valencia, Campione, Weiner, and Bazzi (as cited in Haywood & Lidz, 2007, p. 80) who used an experimental dynamic test approach with several control groups and an overall sample of 196 sixth-grade students. They reported weak but positive effects for their reading program. Post-test results revealed increased strategy use and improved reading comprehension for those students who had been moderately and strongly exposed to scaffolding strategies during the intervention. Moreover, this enhancement in strategy use persisted for at least five months. The more intense the scaffolding procedure during the intervention had been, the higher the retest performance at five-month follow-up.

In a further domain-specific procedure, Tissink, Hamers, and van Luit (1993) applied their learning program to reading and spelling in order to predict future school achievement. The intervention comprised a set of non-standardized prompts (repeated presentation, revelation of item structure, provision of solution strategies and modeling) which were implemented when errors in students' answers occurred. Their procedure is suitable for preschool children and first graders or older low-performing children. Most importantly, the authors found that domain-specific tests were better predictors of school achievement than domain-general tests. Students' learning curves differed according to domain, showing no evidence of a general learning ability. The assessment of a text-comprehension-specific learning ability, as measured by a dynamic test in the domain of reading, should therefore prove highly predictive of future reading comprehension development.

Another study was conducted by Kletzien and Bednar (1990) with fifth-grade students. Initial reading level was established using a reading inventory and students' reading process and strategy utilization by means of probes and observations. The intervention program comprised strategy training including, for example, visualization as well as guided and independent practice sessions. Finally, students were administered a parallel version of the reading inventory. The observed increase in reading level was significantly correlated with reading improvement six months after the assessment. Kletzien and Bednar thus conclude that teachers who instruct students need a “firm understanding of strategies, their use, and ability to infer strategy use from reader responses” as well as the “expertise in utilizing a range of instructional techniques.” (p. 532)

For the construction and application of suitable instructional techniques, a conception of inferential processes and interventions for the development of tasks and feedbacks within dynamic reading-competence tests is required. Dynamic assessments in the domain of reading and text comprehension therefore often rely on instruction and practice in meta-cognitive knowledge (including strategies) which is specific to certain reading tasks and goals. Successful dynamic assessments further depend on the prompting of domain-specific processes which are essential for the fulfillment of task requirements. The major goal of train-within-tests is an efficient diagnostics of reading competence whereas the reported test-train-tests have their focus on competence improvement. To foster learning and understanding in the focused domain, instructions or hints and feedback are used to observe students’ responsiveness to the given support. From a diagnostics point of view this responsiveness allows the examiner to improve interpretations of children’s actual competence range and predictions of further development. Especially train-within-test formats of dynamic assessments underscore the importance of reliable and valid competence diagnostics rather than trainings of competencies.

Conceptualizing a Train-Within-Test of Reading Competence

The development of a dynamic reading-competence test considerably differs from the construction of dynamic assessments in other cognitive domains, as can be illustrated by a comparison with the construction of dynamic tests for reasoning ability: Cognitive components of reasoning tasks are well investigated (Carpenter, Just, & Shell, 1990). For figural reasoning tasks, for example, difficulty is often associated with the number of varying criteria (e.g., shape, color, size etc.) that have to be taken into account. These difficulty-generating task features are used to develop feedback which is provided in the course of dynamic reasoning assessments. In the case of unsuccessful trials, the assessment includes a sequence of feedbacks with increasing complexity and well-defined useful strategies which gradually lead to the correct solution.

In contrast, reading comprehension is a more complex construct involving multi-level processes. In order to comprehend successfully, that is, to gain meaning from written text for a particular purpose, the reader must engage in various processes at the word-, sentence- and text-level. The reader is required to identify a series of letters as a word, access the meaning of words, and integrate individual word meanings or sentence meanings into coherent sentence- and text-level representations (Best, Rowe, Ozuru, & McNamara, 2005). When assessing reading competence dynamically, it seems promising to focus on distinct cognitive processes which are relevant to specific levels of text comprehension that are addressed. In reading assessments, this often includes the generation of inferences, which are known to represent a crucial feature of deep-level understanding. Generating inferences leads to text-based and knowledge-based connections both within and across sentences (Best, et al., 2005). This involves connecting several idea units (propositions) distributed across the text and filling in missing information by activating prior knowledge from long-term memory in an effort to construct global meaning from the text (Artelt, et al., 2005; Artelt, Schiefele, & Schneider, 2001; Graesser, Singer, & Trabasso, 1997). Readers might infer “goals and plans that motivate characters’ actions, character traits, [...], causes of events, the consequences of events and actions, [...], the global theme or point of the text (and) the reference of nouns and pronouns” (Graesser, Millis, & Zwaan, 1997, p. 181). Providing adequate help or feedback within dynamic assessments of reading in a train-within-test design not only has to take into account the specific constraints of the train-within-test format as well as specific cognitive processes of text comprehension, but also general findings from feedback-research.

Hattie and Timperley (2007) define feedback as information provided to a person regarding aspects of his or her performance or understanding. As to the type of feedback a general distinction can be made between verification and elaboration. Verification addresses an answer’s correctness indicating the performance level achieved (e.g., “right – wrong”). This feedback type is the most common form of intervention provided in dynamic tests; certainly due to its simplicity at least in the domain of intelligence (N. Beckmann, Beckmann, & Elliott, in press). In contrast, elaboration provides additional information by means of relevant cues. Elaborated feedback, which offers more variations than verification does, can address the task or topic, particular errors, or responses. A large body of educational research shows that feedback effectiveness varies according to the type of feedback with greatest effects being achieved for elaborated feedback (Bangert-Drowns, Kulik, Kulik, & Morgan, 1991; Kluger & DeNisi, 1996; Shute, 2008). However, Kulhavy and Stock (1989) already argued that effective feedback includes both verification and elaboration.

Due to the test procedure, within train-within-test designs only brief feedback-interventions are possible. For this reason, elaborated feedback needs to be specific as it contains a rather high level of information on how to proceed or why a specific response was incorrect or accurate. When conceptualizing a dynamic test of reading competence that focuses on (causal) inferences concerning processes on shallow (e.g., local coherence) as well as deep levels of comprehension (e.g., global coherence, situational model), elaborated feedback might give that kind of error-specific explanations. That is, the learner is provided with an explanation of why his or her response does not display an accurate inference or causal relationship between several units of the text. Another feedback-intervention might provide solution-oriented prompts. For the sake of this, cognitive aids are applied to elicit a certain inference and to guide the learner to the correct solution without offering too much information. A further intervention in a train-within-test of reading competence could address meta-cognitive processes which are known to be highly relevant in the text comprehension process. Within such interventions, learners might be prompted to reflect on their monitoring performance or evaluating task requirements when responding to the test items.

The development of tasks that refer to high-level inferential processes is challenging because, in contrast to reasoning tasks, deep-level processes of reading comprehension are usually embedded in shallow-level processes. Furthermore, the implementation of a sequence of feedbacks – as found in dynamic reasoning tests – is difficult to realize given that single task-solving steps in reading comprehension are less apparent than is the case in other domains. Nevertheless, successful feedback which is suited to the purposes of dynamic assessment must take such domain-specific complexity into account.

Conclusion

To summarize, the domain of reading competence has rarely been subject to investigations in the framework of dynamic assessments. The studies conducted so far have often implemented reading strategies in terms of training interventions rather than focusing on reliable and valid reading competence measurement. In aiming to develop a conclusive approach for the dynamic assessment of reading competence, findings from different fields of research must be integrated. On the one hand, the various processes involved in reading comprehension provide many opportunities for feedback-interventions which may potentially enhance learning and understanding. On the other hand, elaborated and learning-oriented feedback is known to outperform performance-oriented feedback. We therefore advocate that, as a critical feature of dynamic reading-comprehension assessment, effective feedback should be closely related to targeted comprehension processes, namely inferences. A central problem of a train-within-test of reading competence concerns dealing with performance changes induced by the test procedure itself. IRT models can deal with this problem and provide detailed measures of performance and learning ability in the domain of reading. The potential of a reading competence test in a train-within-test format has to be evaluated in upcoming research.

Acknowledgments:

The preparation of this paper was supported by grant AR 307/7-1 from the German Research Foundation (DFG) in the Priority Program "Models of Competencies for Assessment of Individual Learning Outcomes and the Evaluation of Educational Processes" (SPP 1293).

References

Artelt, C., McElvany, N., Christmann, U., Richter, T., Groeben, N., Köster, J., et al. (2005). *Expertise - Förderung von Lesekompetenz [Fostering Reading Competence]*. Bonn: Bundesministerium für Bildung und Forschung.

Artelt, C., Schiefele, U., & Schneider, W. (2001). Predictors of reading literacy. *European Journal of Psychology of Education, 16*(3), 363-383.

Bangert-Drowns, R. L., Kulik, C., Kulik, J. A., & Morgan, M. T. (1991). The instructional effect of feedback in test-like events. *Review of Educational Research, 61*, 213-238.

Beckmann, J. F. (2001). *Zur Validierung des Konstrukts des intellektuellen Veränderungspotentials [On validation of the construct of intellectual change potential]*. Berlin: logos.

Beckmann, J. F., & Guthke, J. (1999). *Psychodiagnostik des schlußfolgernden Denkens [Assessing reasoning ability]*. Göttingen: Hogrefe.

Beckmann, N., Beckmann, J. F., & Elliott, J. G. (in press). Self-confidence and performance goal orientation interactively predict performance
in a reasoning test with accuracy feedback. *Learning and Individual Differences*.

Berman, J., & Graham, L. (2002). School Counsellor Use of Curriculum-Based Dynamic Assessment. *Australian Journal of Guidance and Counselling, 12*(1), 21-40.

Best, R. M., Rowe, M., Ozuru, Y., & McNamara, D. S. (2005). Deep-level comprehension of science texts. *Top Lang Disorders*, *25*(1), 65-83.

Budoff, M. (1987a). Measures for Assessing Learning Potential. In C. S. Lidz (Ed.), *Dynamic assessment: An interactional approach to evaluating learning potential* (pp. 173-195). New York: Guilford Press.

Budoff, M. (1987b). The validity of learning potential assessment. In C. Lidz (Ed.), *Dynamic assessment: An interactional approach to* *evaluating learning potential* (pp. 53-81). New York: Guilford Press.

Campione, J. C., & Brown, A. L. (1987). Linking dynamic assessment with school achievement. In C. S. Lidz (Ed.), *Dynamic assessment: An international approach evaluating learning potential* (pp. 82-115). New York: Guilford Press.

Campione, J. C., & Brown, A. L. (1990). Guided learning and transfer: Implications for approaches to assessment. In N. Frederiksen, R. Glaser, A. Lesgold & M. G. Shafto (Eds.), *Diagnostic monitoring of skill and knowledge acquisition*. Hillsdale: Erlbaum.

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

Carlson, J. S., & Wiedl, K. H. (1980). Applications of a dynamic testing approach in intelligence assessment: Empirical results and theoretical formulations. *Zeitschrift für Differentielle und Diagnostische Psychologie, 1*, 303-318.

Carlson, J. S., & Wiedl, K. H. (2000). The validity of dynamic assessment. In C. S. Lidz & J. G. Elliott (Eds.), *Dynamic assessment: Prevailing models and applications* (Vol. 6, pp. 681-712). Oxford, UK: Elsevier.

Carpenter, P. A., Just, M. A., & Shell, P. (1990). What One Intelligence Test Measures: A Theoretical Account of the Processing in the Raven Progressive Matrices Test. *Psychological Review, 97*(3), 404-431.

Cook, T. D., & Campbell, D. T. (1976). The design and conduct of quasi-experimentations and true experiments in field settings. In M. Dunnette (Ed.), *Handbook of industrial and organizational psychology*. Skokie, Ill.: Rand McNally.

Desoete, A., Roeyers, H., Buysse, A., & De Clercq, A. (2002). Dynamic assessment of metacognitive skills in young children with mathematics learning disabilities. In G. M. Van der Aalsvoort, W. C. M. Resing & A. J. J. M. Ruijssenaars (Eds.), *Learning potential assessment and cognitive training: Actual research and perspectives in theory building and methodology* (pp. 307-333). Amsterdam: Elsevier Science.

Dillon, R. F. (1997). Dynamic Testing. In R. F. Dillon (Ed.), *Handbook on testing* (pp. 164-186). Westport: Greenwood Press.

Elliott, J. G. (2000). Dynamic assessment in educational contexts: Purpose and promise. In C. S. Lidz & J. G. Elliott (Eds.), *Dynamic assessment: Prevailing models and applications* (Vol. 6, pp. 713-740). Oxford,UK: Elsevier.

Embretson, S. E. (1987). Toward development of a psychometric approach. In C. Lidz (Ed.), *Dynamic assessment: An interactional approach to evaluating learning potential* (pp. 141-170). New York: Guilford Press.

Embretson, S. E. (2000). Multidimensional Measurement from Dynamic Tests: Abstract Reasoning Under Stress. *Multivariate Behavioral Research, 35*(4), 505-542.

Feuerstein, R., Rand, Y., Haywood, H. C., Hoffmann, M., & Jensen, M. R. (1983). *Learning Potential Assessment Device. Manual*. Jerusalem: HWCRI.

Flammer, A. (1975). A model relating transfer of learning and correlation. *Psychometrika, 40*(4), 487-503.

Graesser, A. C., Millis, K. K., & Zwaan, R. A. (1997). Discourse comprehension. *Annual Reviews of Psychology, 48*, 163-189.

Graesser, A. C., Singer, M., & Trabasso, T. (1997). Constructing inferences during narrative text comprehension. *Psychological Review, 101*, 371-395.

Guthke, J. (1982). The learning test concept - an alternative to the traditional static intelligence test. *The German Journal of Psychology, 6*, 306-324.

Guthke, J., Beckmann, J. F., & Wiedl, K. H. (2003). Dynamik im dynamischen Testen [Dynamics in dynamic testing]. *Psychologische Rundschau, 54*(4), 225-232.

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

Guthke, J., & Wiedl, K. H. (1996). *Dynamisches Testen: Zur Psychodiagnostik der intraindividuellen Variabilität [Dynamic Testing]*. Göttingen: Hogrefe.

Hattie, J., & Timperley, H. (2007). The power of feedback. *Review of educational research, 77*(1), 81-112.

Haywood, H. C., & Lidz, C. S. (2007). *Dynamic Assessment in Practice: Clinical and Educational Applications*. New York: Cambridge University Press.

Hessels, M. G. P. (2000). The learning potential test for ethnic minorities (LEM): 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* (Vol. 6, pp. 109-131). Oxford,UK: Elsevier.

Hessels, M. G. P. (in press). Estimation of the predictive validity of the HART by means of a dynamic test of geography. *Journal of Cognitive Education and Psychology, 8*.

Kintsch, W. (1998). *Comprehension: A paradigm for cognition*. Cambridge: Cambridge University Press.

Klauer, K. C., Kauf, H., & Sydow, H. (1994). Experimentelle Validierung eines Lernmodells für Kurzzeitlerntests [Experimental validation of a model for short-time learning tests]. *Diagnostica, 40*(2), 124-142.

Klauer, K. C., & Sydow, H. (1992). Interindividuelle Unterschiede in der Lernfähigkeit. Zur Analyse von Lernprozessen bei Kurzzeitlerntests [Interindividual differences in learning ability: Analysis of learning with short-time learning tests]. *Zeitschrift für Differentielle und Diagnostische Psychologie*, *13*(3), 175-189.

Kletzien, S. B., & Bednar, M. R. (1990). Dynamic Assessment for At-Risk Readers. *Journal of Reading*, *33*(7), 528-533.

Kluger, A. N., & DeNisi, A. (1996). Effects of feedback intervention on performance: A historical review, a meta-analysis, and a preliminary feedback intervention theory. *Psychological Bulletin*, *119*(2), 254-284.

Kozulin, S., & Garb, E. (2002). Dynamic assessment of EFL text comprehension. *School Psychology International*, *23*(1), 112-127.

Kulhavy, R. W., & Stock, W. A. (1989). Feedback in written instruction: The place of response certitude. *Educational Psychology Review, 1*(4), 279-308.

Lussier, C. M., & Swanson, H. L. (2005). Dynamic assessment: a selective synthesis of the experimental literature. In G. M. van der Aalsvoort, W. C. M. Resing & A. J. J. M. Ruijssenaars (Eds.), *Learning potential assessment and cognitive training: actual research and perspectives in theory building and methodology* (pp. 65-87). New York: Elsevier.

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

Palincsar, A. S., Brown, A. L., & Campione, J. C. (1991). Dynamic Assessment. In H. L. Swanson (Ed.), *Handbook on the assessment of learning disabilities: Theory, research, and practice* (pp. 75-94). Austin, Texas: Pro-ed.

Poehner, M. E. (2008). *Dynamic Assessment: A Vygotskian Approach to Understanding and Promoting L2 Development*. US: Springer.

Rost, J. (2004). *Lehrbuch Testtheorie - Testkonstruktion [Test theory and test development]*. Bern: Huber.

Schneider, E., & Ganshow, L. (2000). Dynamic assessment and instructional strategies for learners who struggle to learn foreign language. *Dyslexia*, *6*, 72-82.

Shute, V. J. (2008). Focus on Formative Feedback. *Review of Educational Research, 78*(1), 153-189.

Sijtsma, K. (1993a). Classical and modern test theory with an eye toward the learning potential testing. In A. J. J. M. Ruijssenaars, J. H. M. Hamers & K. Sijtsma (Eds.), *Learning potential assessment: Theoretical, methodological and practical issues* (pp. 117-133). Lisse, Netherlands: Swets & Zeitlinger Publishers.

Sijtsma, K. (1993b). Psychometric issues in learning potential assessment. In A. J. J. M. Ruijssenaars, J. H. M. Hamers & K. Sijtsma (Eds.), *Learning potential assessment: Theoretical, methodological and practical issues* (pp. 175-185). Lisse, Netherlands: Swets & Zeitlinger Publishers.

Tissink, J., Hamers, J. H. M., & van Luit, J. E. H. (1993). Learning potential tests with domain-general and domain-specific tasks. In A. J. J. M. Ruijssenaars, J. H. M. Hamers & K. Sijtsma (Eds.), *Learning potential assessment: Theoretical, methodological and practical issues* (pp. 243-266). Lisse, Netherlands: Swets & Zeitlinger Publishers.

Vygotsky, L. S. (1964). *Denken und Sprechen [Thought and language]*. Berlin: Akademie-Verlag.