



Universiteit
Leiden
The Netherlands

Computerised dynamic testing: An assessment approach that tailors to children's instructional needs

Touw, K.W.J.

Citation

Touw, K. W. J. (2020, September 17). *Computerised dynamic testing: An assessment approach that tailors to children's instructional needs*. Retrieved from <https://hdl.handle.net/1887/136755>

Version: Publisher's Version

License: [Licence agreement concerning inclusion of doctoral thesis in the Institutional Repository of the University of Leiden](#)

Downloaded from: <https://hdl.handle.net/1887/136755>

Note: To cite this publication please use the final published version (if applicable).

Cover Page



Universiteit Leiden



The handle <http://hdl.handle.net/1887/136755> holds various files of this Leiden University dissertation.

Author: Touw, K.W.J.

Title: Computerised Dynamic Testing: An assessment approach that tailors to children's instructional needs

Issue Date: 2020-09-17

Chapter 6

GENERAL DISCUSSION



Within the framework of inclusive education, teachers are increasingly being asked to cater to their students' individual educational needs. Process-oriented assessment techniques, such as dynamic testing, can be of much value in this regard. Dynamic testing, using a test-training-test format, allows for more tailored measurement of children's level of cognitive ability and has the potential to provide in-depth information about children's learning processes (Elliott, Resing, & Beckman, 2018). Moreover, implementing a dynamic test on a tablet can make the assessment procedure more efficient. Educational and school psychologists have increasingly less time for individual students, which is why the time-efficient and highly structured nature of computerised dynamic testing is a desirable and necessary step in education (Resing, Touw, Veerbeek, & Elliott, 2017). In this dissertation, the effects of computerised dynamic testing were examined in relation to the children's school performances (Chapter 2 and 5), the children's verbal explanations of their task solutions (Chapter 3), the individual differences in the children's needs for prompts during training (Chapter 4 and 5), the factors that influence individual differences in solving the tasks (Chapter 3), and the practical value of such a testing method for teachers (Chapter 5). In addition, the potential value of computerised dynamic testing was extended to a group of children with a developmental language disorder who attended special education in the Netherlands (Chapter 4).

In the current chapter, a summary of the key findings of the studies in this dissertation was provided, followed by some limitations of these studies and the implications for future research. Finally, the value of dynamic testing for educational practice was considered.

Summary of findings

The effect of computerised dynamic testing. The results of the studies involving the dynamic test using a series-completion task on a tablet repeatedly demonstrated that children who received graduated prompts training showed more progression in series completion than children who received no training (Chapter 2, 3 and 4). Children who were dynamically tested showed steeper progression lines from pre-test to post-test compared to children who were statically tested. In the study reported in Chapter 5, a visual analogy task was used. In this study, the children were divided into weak and strong functioning groups, as indicated by the teacher, and all children received training. The study findings revealed that children categorised in the weak functioning group had lower accuracy in analogical reasoning than those in the strong functioning group. However, both groups seemed to benefit from training to a similar extent, indicating that these groups might have differed in their initial level of inductive reasoning, but not in their improvement during dynamic testing. Overall, this dissertation's findings concerning the effectiveness of dynamic testing with two different inductive reasoning tasks mirror and further support a myriad of earlier studies in which the use of dynamic testing principles had an additional effect on children's improvement in inductive reasoning accuracy (e.g., Resing & Elliott, 2011; Resing, Steijn, Xenidou-Dervou, Stevenson, & Elliott, 2011; Resing, Touw

et al., 2017; Stevenson, Hickendorff, Resing, Heiser, & De Boeck, 2013; Tzuriel & Egozi, 2010; Veerbeek, Vogelaar, Verhaegh, & Resing, 2019).

Verbal explanations of children's solutions. The results presented in Chapter 3, concerning whether trained children would use more advanced explanations of their solutions than their untrained peers, revealed that children were categorised most often as using non-responsive and partial-inductive verbal explanations. While the results did not show that training produced different strategy paths for these two categories, significant effects were found for the non-inductive and full-inductive verbal explanations, which the children used less frequently. At the post-test, trained children were found to use fewer non-inductive verbal explanations, and more full-inductive verbal strategies. Although the effects found in this study were small, they lend support to previous findings that children's verbal explanations of their solutions of the series-completion task-items progressed to a more advanced level of reasoning after dynamic training, when compared with their non-trained peers (Resing, Bakker, Pronk, & Elliott, 2017; Resing, Touw et al., 2017; Veerbeek, Verhaegh, Elliott, & Resing, 2017; Veerbeek et al., 2019).

Individual differences in children's need for prompts. In addition, the findings of the studies that are part of this dissertation showed that more in-depth information regarding children's instructional needs can be derived from dynamic test results. The results presented in Chapter 4 and 5 indicated that both typically developing children and children with a developmental language disorder attending special education showed a reduction in the number of prompts they needed in the second training session compared to the first. It appeared that the children learned what they had to focus on when solving the inductive reasoning tasks, approaching the tasks more independently. Overall, large individual differences were found in relation to the instructions they needed during training, supporting previous findings that children display individual differences in their need for instructions, both with regard to the amount and the type of help they needed to learn effectively within their individual zone of proximal development (e.g., Resing, 2013; Resing, Touw et al., 2017).

Factors that influence individual differences. The results concerning inter-individual differences (Chapter 3) revealed that improvement in reasoning accuracy could be predicted most strongly by having received training as well as children's initial reasoning ability: Children who were trained, specifically those with lower initial reasoning abilities, were found to obtain higher scores at post-test than those who were not trained. Moreover, trained children who had higher initial reasoning abilities, and a relatively weaker auditory working memory demonstrated larger improvements from pre-test to post-test than those who had a stronger auditory working memory. These results underline the importance of dynamic testing for unveiling the potential of children with weaker initial inductive reasoning abilities.

Dynamic testing of children with a developmental language disorder. This dissertation also provided a first study into the use of computerised dynamic testing for children with a developmental language disorder in special education (Chapter 4). The results showed that children who had received graduated prompts training improved more in accurately applied transformations than those who were not trained. In addition, it was found that children with lower initial scores showed larger improvements after training than those with higher initial scores, although this finding did not reach significance, possibly because of the small number of participants. Lastly, as mentioned above, at the group level children with a developmental language disorder were found to require fewer prompts in the second than in the first training session, but, at the individual level, showed large individual differences in the number and the type of prompts they needed to solve the inductive reasoning tasks. These findings underline the idea that dynamic testing can serve as a means to gain insight into the progression paths of children whose abilities might be underestimated by static assessment tools.

Dynamic testing and the relationship with children's school performance. This dissertation also investigated the relationship between dynamic and static test outcomes, and children's scholastic performance. School performance was measured by means of teacher judgments, and by standardised school achievement scores as part of the student tracking system (Chapter 2 and 5). In the study presented in Chapter 2, using the series completion task, it was found that the number of prompts children needed during training could significantly predict children's test scores for mathematics and technical reading. In the study reported in Chapter 5, using the analogical reasoning task, it was found that both static and dynamic test outcomes were significantly related to the test scores for mathematics and spelling, specifically for those who were categorised as strong functioning children. Moreover, the number of prompts children required was found to be related most strongly to their performance on the mathematics test, whereas only the weak functioning group's ability to reason by analogy was related to their technical reading performance.

All in all, in accordance with previous studies, these findings suggested that dynamic measures have additional value in predicting school performance, underlining the importance of measuring children's abilities dynamically (e.g., Stevenson, Bergwerff, Heiser, & Resing, 2014).

Practical value for teachers. In Chapter 5, it was also investigated whether information derived from dynamic test results had practical value for teachers. In doing so, teachers and educational consultants were asked what their views were on what dynamic test results could contribute to individual education plans. Both teachers and internal educational supervisors indicated that they evaluated the practical recommendations provided in the dynamic test reports as highly valuable. These findings lend support to suggestions from researchers that the main advantages of dynamic testing include the applicability of the practical recommendations resulting from training (Elliott, 2003; Lauchlan & Carrigan, 2013; Lauchlan & Elliott, 2001).

Limitations and future research

Item construction of the series completion task. The items for the series completion task were carefully constructed based on the series completion model of Simon and Kotovsky (1963). The results of the study described in Chapter 2 indicated that the theoretically expected and empirically retrieved item difficulty levels were associated strongly, while the reliability, as measured by the internal consistency, of the series completion test was modest. A careful analysis of the various items' difficulty levels led to the conclusion that when attempting to solve the task items the children experienced the items as having relatively high or relatively low difficulty, but almost no moderate difficulty. It must be noted that the items' difficulty level, which was relatively high in order to avoid a possible ceiling effect, possibly influenced negatively some of the study results. The use of additional items of moderate difficulty level should be further explored, as this approach could possibly increase the effect of training, in particular for children in the tested age group.

A possible explanation for the lack of items that have a moderate difficulty level could be that children needed to make a change in their solving strategies when dealing with more difficult items. In order to solve the relatively difficult items, they had to focus on more pieces of information, and find more than one solving rule. These potential changes in their use of strategies may be dependent on cognitive flexibility, which, as research shows, is important in children's improvements in inductive reasoning tasks (Stad, Wiedl, & Resing, 2016). The potential role of children's executive functioning levels, such as cognitive flexibility and working memory, in how they perform on the series completion task used in this study should therefore also be taken into account in future research.

Verbal aspect of the dynamic test. Another limitation of this study concerned the verbal aspect of the instructions and the responses when the children were asked to provide explanations of their solutions. First, the children were asked for verbal explanations to gain more information about their strategy use. However, for younger children verbalisations might not always be a reliable indicator of the processes in problem-solving (Ericsson & Simon, 1984; Resing, Xenidou-Dervou et al., 2012). Therefore, it is recommended that future studies into strategy use during dynamic testing focuses on both children's verbal explanations and their problem-solving behaviours, as analysing these behaviours potentially provides a more profound insight into the nature of individual differences.

Second, although the items of the series completion task consisted of geometric figures, and the pre and post-test phases involved a minimal level of spoken language, which makes this approach, in principle, an appropriate assessment tool for children of diverse cultural and linguistic backgrounds, feedback provided during training may have relied more on spoken language than on visual processing. While this study sought to develop a testing procedure utilising visual effects to convey instructions, spoken language was still used to provide the prompts' content. In previous research, it was suggested that relying on spoken language in testing situations can be complicated for children that have

language-related difficulties (e.g., Camilleri & Botting, 2013). Future studies that include children with such language-related problems may need to adapt the verbal instructions provided further to make the training more appropriate and effective.

Extension of the technical possibilities. Future research should also investigate more thoroughly the technological benefits of tablet-administered dynamic tests, as usage of computer-assisted technology can yield novel methods for fine-grained measuring of problem-solving processes (Resing & Elliott, 2011; Resing, Touw et al., 2017; Tzuriel & Shamir, 2002). Moreover, computerised testing can generate large quantities of data, which can potentially be valuable information for further understanding and improving individualised instructions in the classroom, the processes occurring during problem-solving, and didactic interventions (Passig, Tzuriel, & Eshel-Kedmi, 2016; Resing & Elliott, 2011; Stevenson, Touw et al., 2011). Moreover, computerisation of instructions and prompting facilitates adaptive, yet standardised, testing procedures, providing testees with only those instructions they need in order to complete the test items successfully (Resing, Touw et al., 2017). For this reason, future studies should focus on the benefits of dynamic tests that are adaptive, and administer items of which the difficulty level lays as close as possible within individual children's zones of proximal development, utilising the series completion task items tailored to the needs of larger age groups.

Domain-specific vs. domain-general skills. Another focus of future research concerns the relationship between, on the one hand, domain-specific and domain-general skills, and, on the other, academic performance. In two studies performed for this dissertation, relationships were found between two inductive reasoning tasks, which are considered to involve domain-general skills, and various scholastic tasks such as mathematics, reading and spelling, which are considered to involve domain-specific skills. Such research would be in line with a current debate on the role that both domain-specific and domain-general factors play in academic performance of children (e.g., Cromley, Snyder-Hogan, & Luciw-Dubas, 2010; Geary, 2011; Passolunghi & Lanfranchi, 2012). School curricula should not predominantly focus on the knowledge and skills underlying domain-specific skills, such as reading, writing and arithmetic, but also include domain-general reasoning skills, for example in relation to categorising information processing. This would enable children to integrate different stimuli into meaningful representations, and determine whether certain stimuli are relevant for completing a task. Given the importance of both domain-specific and domain-general skills in academic performance, more research is needed into computerised dynamic testing in different domains. In such research, the principles behind domain-general computerised dynamic tests could be linked to specific school subjects, extending the findings and insights from the current dissertation. Such studies should also have a special focus on translating the testing outcomes to teachers, so they can use them

in their educational practice, for instance when drawing up educational plans (e.g., Bosma, Stevenson, & Resing, 2017).

Practical value for education

When a teacher has concerns about a child's school performance, the child may be referred for assessment. Most often questions are raised about the child's intelligence level. A common follow-up question is how the teacher can help this child in the classroom. Typically, the child's intelligence level will be assessed with a static intelligence test. However, as noted above, such tests can underestimate the child's cognitive abilities. Moreover, only limited practical information about a child's educational and instructional needs can be derived from a static test. Gipps (1994) states that the primary purpose of assessment should be to support the teaching and learning process, whereas Lidz (1992) has argued that the main aim of assessment should be more practical, with a focus on answering questions such as 'how do we teach this child?' and 'how can we help teachers individualise their programmes?' (Lidz, 1992, p. 207). The more process-oriented assessment approach of dynamic testing is a necessary and valuable addition in this regard. Since it yields more fine-grained information, which cannot be easily derived from a static testing situation, dynamic testing can be used as hands-on starting points for drawing up didactic interventions and strategies (e.g., Jeltova et al., 2007). For example, some children seem to have benefited greatly from general, metacognitive instructions, while others have been shown to benefit more from item-specific step-by-step scaffolding. Such fine-grained information about children could help teachers in providing differentiated instruction in the classroom.

The results of this dissertation suggest that teachers' estimates of children's cognitive-intellectual ability and ability to learn are predominantly based on their scholastic performance, and, to a minor extent on their needs for instructions, use of strategies or other individual learner characteristics. According to the "Pygmalion effect", teachers' expectations affect students' academic performance (Rosenthal, 1987, 2002; Wang & Cai, 2016). Through dynamic testing, which looks beyond the students' current performance on school tests, teachers' expectations can be raised because more information can be provided concerning children's cognitive growth. By utilising dynamic testing, teachers will be provided with a more in-depth insight into individual children's preferences for instructional needs, and, in doing so, can discover that these needs are not necessarily associated with children's intelligence levels. Individual differences in instructional needs are to be found amongst children of all ranges of intelligence (Vogelaar & Resing, 2016). Two children who are categorised by their teachers as either low or high functioning in the classroom could have similar instructional needs. Both children could, for instance, benefit markedly, and show substantial improvement, from the metacognitive prompt 'look first, then think carefully, and then provide an answer'. Further adaptation to individual children's instructional needs by tailoring the content of the

prompts to individual educational needs seems to make it more likely that children are educated within their individualised zone of proximal development (Vygotsky, 1962, 1978).

In conclusion, dynamic testing enables provision of information about children's progression paths, which provides schools with a more insightful and in-depth understanding of the factors and processes underlying learning, and contributing to cognitive growth. By following these dynamic testing methods, schools will have the possibility to change the manner in which they provide instructions from a standardised approach to one that includes adaptive teaching and giving of instructions, tailored to the needs of children from all backgrounds.