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

Touw, K.W.J.

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 1

General Introduction

Testing the cognitive abilities of children in clinical and educational settings is most frequently done with static assessment tools. Static tests, such as intelligence tests or school tests often measure the skills and knowledge a child has mastered or acquired at the time of testing. Critics observe, however, that such test outcomes are not an optimal reflection of children's potential for cognitive growth or their ability to learn (Elliott, Resing, & Beckmann, 2018; Haywood & Lidz, 2007; Resing, Elliott, & Vogelaar, in press), providing limited information about processes that underlie children's learning (Grigorenko, 2009; Jeltova et al., 2007). In addition, research has shown that conventional static tests can be biased towards children that have not had the chance to demonstrate optimal performance in learning, such as those with disadvantaged backgrounds (Tzuriel, 2000), culturally diverse backgrounds (Gillam & Peña, 2004), language deficits (Martin, 2015), or special needs (Peltenburg, Van den Heuvel-Panhuizen, & Doig, 2009).The ways in which individual children learn and perform in a school context differ significantly, and can result in different instructional needs which cannot easily be abstracted from static test results. In response, the interest in assessment techniques that are more process oriented, such as dynamic testing, has grown.

Measuring children's ability to learn by means of dynamic testing principles constitutes a different perspective. As opposed to static tests, in which the focus is predominantly on what children can achieve independently, without help, dynamic tests focus on how a child can perform after training, and on the instructions they need during training. Consequently, dynamic testing not only allows for comparing a child's performance with their peers, but, more importantly, with his or her own performance, taking into account a child's past and current performance on a task. In addition, dynamic testing yields information about the specific forms of help that may be beneficial for an individual child (Bosma & Resing, 2012; Resing & Elliott, 2011). As the ability to learn is an important part of dynamic testing, this form of testing allows for measuring the different processes that play a role in learning (Resing, 2013). Such information can be useful, for example, for educational psychologists seeking to inform teachers about additional forms of intervention in the classroom (Grigorenko, 2009; Jeltova et al., 2011). This approach is in line with the growing responsibility of schools to provide education that suits all children, including children with developmental delays or learning disabilities (Kurth, Miller, Toews, Thompson, Cortés et al., 2018; Pameijer, 2016). The policy of many countries around the world is to move towards inclusive education, which emphasises that children and young people with special educational needs (SEN) and disabilities have the right to be included in mainstream education (Jeltova et al., 2011). In this endeavour, it is important to obtain a more profound insight into the different mediating and moderating factors facilitating an optimal learning environment for all children, including those with disadvantaged learning opportunities in the past or disabilities.

Another focus in education that is becoming more important rapidly is the use of technology to assist teachers' teaching. The increased importance of using technology in the classroom highlights the potential of computerised dynamic testing as a promising instrument for providing teachers and other educational professionals with insight into individual children's ability and educational needs.

The main aims of this dissertation were to provide insight into the advantages of dynamic testing, examine the possibilities of using computerised testing to increase efficiency in the assessment of children's abilities, and provide more information on the practical value of test outcomes for education. This dissertation includes a series of studies investigating the additional effects of dynamic testing beyond those of (repeated) static testing on an inductive reasoning task performed on a tablet, for both normally developing children and children with a developmental language disorder. The focus of this dissertation includes the possible factors involved in the progression of the children's series completion abilities, the relationship between dynamic test results and school achievement, children's need for instruction and the advantages of the dynamic test results for teachers.

This chapter shortly introduces the theory of dynamic testing and the development of the computerised test used for conducting this research. This theoretical background is followed by a section on the dynamic testing of a specific subgroup: Children with a developmental language disorder. The chapter concludes with an outline of the studies included in this dissertation.

Dynamic Testing

The theoretical underpinnings of dynamic testing can, for a large part, be found in the theory of the zone of proximal development (ZPD), as posed by Vygotsky (1978), which has been a major source of influence for educational practice (Elliott et al., 2018). A child's ZPD can be defined as the distance between a child's level of independent performance (without help), and a child's performance with (more capable) others. Vygotsky theorised that children differ in their zone of proximal development and in their need for help in this interaction with adults or peers. Based on this theory, dynamic testing incorporates feedback and training into the assessment procedures, allowing changes that occur during this assessment to be examined and more information about the individual's ZPD to be provided (Elliott et al., 2018; Grigorenko, 2009; Haywood & Lidz, 2007; Resing, 2013; Resing & Elliott, 2011; Tzuriel, 2011). The format of testing and training differs among types of dynamic testing (Sternberg & Grigorenko, 2002), but the use of individualised instruction and feedback is a common element in all dynamic testing procedures.

One format of dynamic testing described in the literature is the 'sandwich format', consisting of a pre-test, which the child completes independently, followed by a training, and a post-test, which the child also completes independently. A second dynamic testing format is the so-called 'cake format', which entails providing help immediately when a child demonstrates a difficulty in coming to a correct solution (Elliott, 2003). In the studies conducted for this dissertation, the sandwich format was used, which facilitated a high level of standardisation of the test and training procedures, including the feedback provided. Dynamic tests with a sandwich format can have different instruction procedures during the assessment phases. One approach, in which standard static testing is combined with standardised feedback during the training process, is the graduated prompts approach (e.g., Campione & Brown, 1987; Ferrara, Brown, & Campione, 1986; Resing, 1997; Resing & Elliott, 2011). This procedure refers to a graduated, hierarchical method for training children between test phases. When the task cannot be solved independently, the child is gradually prompted towards the correct solution with predetermined prompts. This approach offers a hierarchy of prompts ranging from a general metacognitive level to those at a more concrete task-specific level. Moreover, the prompts provided are restricted: the child is provided with only the minimum number of prompts necessary to accurately solve the task.

Research has repeatedly revealed that dynamic testing has the potential to obtain insight into weaknesses and strengths in individual children's cognitive abilities (e.g., Bosma, Stevenson, & Resing, 2017; Hill, 2015; Jeltova et al., 2007; Resing, Bakker, Pronk, & Elliott, 2017; Resing & Elliott, 2011; Resing et al., in press; Tzuriel, 2000). Dynamic testing aims to enable in-depth analysis of the underlying learning and cognitive processes potentially occurring during testing (Resing, Xenidou-Dervou, Steijn, & Elliott, 2012; Veerbeek, Vogelaar, Verhaegh, & Resing, 2019). Consequently, results derived from dynamic testing may provide useful information on which didactic interventions can be based (Jeltova, et al., 2007). However, in spite of the potential advantages of dynamic testing in obtaining insight into individualised instruction, these tests are used sparsely in practice (Elliott & Resing, 2015; Elliott et al., 2018; Stringer, 2018), because, among other reasons, administration of a dynamic test often takes more time than is needed for conventional static testing methods (i.e. 'extra time' relative to other testing approaches). This dissertation attempted to overcome this limitation by developing a computerised dynamic test which can be administered easily on a tablet.

Computerised assessment

The use of technology in school settings has increased significantly in recent decades (De Vries, 2018; Domingo & Garganté, 2016). Teachers frequently use, for example, digital whiteboards and tablets to assist their teaching. Because the use of tablets is already implemented in the school environment, this technological instrument was used to administer the dynamic test in most of the studies included in the dissertation. It was assumed that using a tablet would allow for standardisation of the assessment instrument, maintaining the option of flexibility, and for test outcomes to be interpreted more efficiently and with more ease, facilitating evaluation of these outcomes with teachers. In various studies, the usefulness of computerised instruction during a dynamic test has been examined (e.g., De Beer, 2013; Poehner & Lantolf, 2013; Resing & Elliott, 2011; Stevenson, Touw, & Resing, 2011; Tzuriel & Shamir, 2002; Wu, Kuo, & Wang, 2017). For example, Stevenson, Touw et al.

(2011) analysed differences between a computer-assisted and paper-and-pencil dynamic test of analogical reasoning, and found that these test versions had an equivalent difficulty level, while administration of the computer-assisted version took less time. Computerised instruments, furthermore, have the potential to generate useful information for individualisation of instruction and intervention in educational practice, such as information about the processes occurring during problem solving (Resing & Elliott, 2011). Examples of such information include completion time and the different steps children take when solving problems (Resing, Steijn, Xenidou-Dervou, Stevenson, & Elliott, 2011; Resing, Touw, Veerbeek & Elliott, 2017).

Dynamic testing researchers developed a series completion test using an electronic console based on sensor technology in combination with electronic tangibles (e.g., Resing & Elliott, 2011; Resing et al., 2011; Resing, Touw et al., 2017). These studies involving the electronic series completion 'puppet' task showed that children demonstrated improvements in performance as a result of dynamic testing, including greater accuracy and more advanced strategy use, compared to untrained children, who received a static test. Despite these positive and promising results, the practical implications of using an electronic console presents difficulties in clinical and educational settings due to the costs and administration time. For this reason, the possibility to develop a series completion test using electronic technology that can be more easily incorporated into the classroom was studied in this dissertation. The test was administered on a tablet instead of on an electronic console because many teachers already make use of a tablet in their teaching. Moreover, this series completion test used a whole-figure geometric solution instead of a multiple-parts item solution, which was supposed to decrease administration time.

Inductive reasoning

Dynamic tests often consist of tasks, such as series completion or solving analogies, that involve inductive reasoning, a form of high-level cognition that plays a key role in cognitive and intellectual abilities (Goswami, 2002, Klauer & Phye, 2008; Perret, 2015). The ability to solve novel problems without applying prior knowledge, also known as fluid reasoning, is one of the broad abilities of intelligence described by the Cattell–Horn–Carroll theory (e.g., Cattell, 1963; McGrew & Wending, 2010). Inductive reasoning tasks, including analogical reasoning, categorisation and seriation, all require rule-finding processes that search for underlying similarities and differences between task features or in the relations between these features (e.g., Hayes, Stephens, Ngo, & Dunn, 2018; Klauer, 2014; Perret, 2015). Both cognitive and metacognitive processes are involved in making these comparisons: All task characteristics should be examined, and relevant similarities, differences, and relationships should be identified. Subsequently, all these features should be taken into consideration when determining how to solve a problem, as a child should evaluate whether he or she has already solved such problems, decide on the order in which the problem should be solved, and monitor whether the solution they arrived at is correct. Inductive reasoning is also considered critical for scholastic learning (Csapó, 1997; Goswami, 2002; Primi, Ferrão, & Almeida, 2010), as reading, writing, and arithmetic are considered dependents of, among other variables, reasoning skills (De Koning & Hamers, 1999; Peng, Wang, Wang, & Lin, 2019). The domain-independent applicability of inductive reasoning and the fundamental role of this process in child development have also been widely recognised (De Koning, Hamers, Sijtsma & Vermeer, 2002; Molnár, Greiff, & Csapó, 2013; Resing, 2000).

With regard to series completion tasks, studies have shown that already at the age of three, young children have a basic understanding of seriation, and can solve seriation tasks in which the end position of a series should be identified (Siegel, 1972). One aim of this dissertation was to examine whether and how 6–8 year-old children solve inductive reasoning tasks because the early school years are considered to be a period of peak improvement in reasoning ability (e.g., Siegler & Svetina, 2002) and have a profound impact on achievement in school (Wendelken et al., 2017).

Dynamic testing of children with a language disorder

Dynamic testing may be a more accurate and sensitive method for measuring children's cognitive abilities and growth, especially for groups of children whose cognitive abilities may be underestimated by traditional static testing instruments and methods, such as children with language problems (Martin, 2015). Therefore, another aim of this dissertation was to learn more about the possible positive effects of dynamic testing for children with a language disorder. In this study, the term 'developmental language disorder' is used instead of the term 'specific language impairment'. In 2016–2017, two linked projects were conducted to help develop a consensus on the criteria and terminology used for children's language disorders (Bishop, Snowling, Thompson, & Greenhalgh, 2016; Bishop et al., 2017). One outcome was the recommendation that the terminological use of specific language impairment be abandoned in favour of developmental language disorder. In the Netherlands the term developmental language disorder is also used (*Taalontwikkelingsstoornis,* in Dutch).

The research concerning the value of dynamic tests for assessing the inductive reasoning abilities of children with a language disorder is limited. A few studies have been conducted in the field of dynamic testing and language impairments or disorders (e.g. Camilleri & Botting, 2013; Glaspey & Stoel-Gammon, 2007; Hasson & Botting, 2010; Hasson, Dodd, & Botting, 2012; Kapantzoglou, Restrepo, & Thompson, 2012; Martin, 2015). These studies mostly focused on dynamic assessment of language skills, such as receptive vocabulary, expressive grammar, syntax and phonology, and showed that a dynamic approach provides both quantitative and qualitative information which can be used for designing more individualised language interventions. Although several studies investigating young children's language and cognitive development have shown that language may play an important role in various forms of reasoning and complex problem solving (Hermer-Vazquez, Spelke, & Katsnelson, 1999; Leroy, Parisse, & Maillart, 2012), no studies have focused on the dynamic testing of inductive reasoning with children with a developmental language disorder. Using non-verbal inductive reasoning tasks, such as those used in the studies for this dissertation, appears to be suitable for children experiencing language difficulties. In the tablet-administered series completion test used in the current dissertation, instructions and feedback were incorporated utilising visual, oral, as well as auditory modalities.

Outline of this dissertation

The main research questions of the studies described in Chapter 2 concerned (1) the potential effects of a computerised dynamic test on children's progression in series completion and (2) the extent to which the outcomes of the dynamic test were related to children's school performance and their teachers' ratings of their school performance and cognitive capabilities. In this study, the focus was on children's improvement in series completion after they received a graduated prompts training, and on the relationship between the outcomes of this dynamic testing procedure and children's performance on static school tests. As opposed to previous studies in the field of dynamic testing, this research used a computerised dynamic test: Instructions and feedback were provided by the tablet instead of the examiner. The use of electronic technology and dynamic testing was combined to overcome the limitations of conventional static testing, and adapt more closely to the children's individual needs.

The study provided in Chapter 3 also described the effects of the computerised dynamic series completion test, but used gain scores based on Item Response Theory (IRT) Rasch scaling as measures for children's potential for learning. The central research question about children's problem solving on a seriation task was extended by also examining the verbal explanations of their solutions and the factors influencing individual differences in solving series-completion tasks, focusing on the potential roles that initial ability (pre-test score on the series completion test), condition, gender, age and working memory play in dynamic test outcomes.

The study reported in Chapter 4 sought to extend earlier findings in dynamic testing with typically developing children to a group of children with language difficulties. The main aim of this was to determine whether dynamically versus statically tested children diagnosed with a developmental language disorder and attending special education showed differential patterns in their serial reasoning performance and their need for instruction. The study examined the effect of graduated prompts training on the children's solving of series completion task-items, considering the number of correctly applied transformations the children utilised in their answer, and then dividing them into a relatively weak and a strong scoring group. The children's instructional needs patterns were examined by investigating potential individual differences in their need for prompts during training.

The main research questions in the study detailed in Chapter 5 focused on the applicability and advances of dynamic testing results for educational practice. This study aimed to provide insight into the potential for learning of first grade children in weak and strong functioning groups, as indicated by the teacher. This study also sought to provide teachers with dynamic test results to enable individualised educational practices for the participating children. The practical value of dynamic testing was examined by asking teachers and educational consultants about the extent to which the dynamic test outcomes achieved in this study could contribute to instruction plans.

Chapter 6 concluded with an overview of the results of the studies and described implications of the key findings for assessment and education, including directions for future research concerning dynamic testing.