



Universiteit
Leiden
The Netherlands

Hidden treasures: Uncovering task solving processes in dynamic testing

Veerbeek, J.

Citation

Veerbeek, J. (2019, April 11). *Hidden treasures: Uncovering task solving processes in dynamic testing*. Retrieved from <https://hdl.handle.net/1887/71235>

Version: Not Applicable (or Unknown)

License: [Leiden University Non-exclusive license](#)

Downloaded from: <https://hdl.handle.net/1887/71235>

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

Cover Page



Universiteit Leiden



The handle <http://hdl.handle.net/1887/71235> holds various files of this Leiden University dissertation.

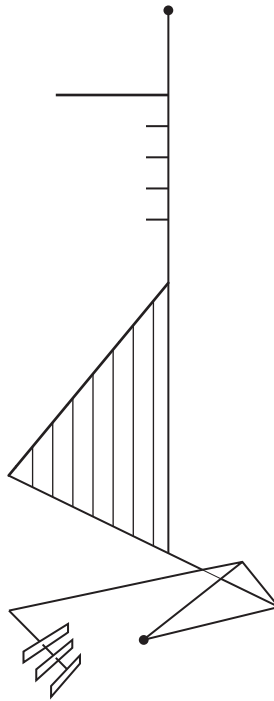
Author: Veerbeek, J.

Title: Hidden treasures: Uncovering task solving processes in dynamic testing

Issue Date: 2019-04-11

CHAPTER 1

General introduction



A large part of children's cognitive development takes place in an educational context. Conventional, static, cognitive ability tests are often used as assessment tools when questions arise regarding children's cognitive capabilities or when they do not show the progression in academic achievements their peers do (Elliott, Resing, & Beckmann, 2018). However, these static assessment instruments are said to provide rather limited product-oriented information regarding the potential reasons behind children's success or failure on the test (Elliott, 2000; Elliott, Grigorenko, & Resing, 2010; Haywood & Lidz, 2007; Sternberg & Grigorenko, 2002). They hardly provide information on individual differences regarding the processes children employ to solve (complex) cognitive tasks (Richard & Zamani, 2003). To address these drawbacks, dynamic testing has been developed, which aims to make an estimate of a child's potential for learning on the basis of feedback or intervention methods that lead to successful task solving behavior. Some forms of dynamic testing focus on process assessment, aiming to obtain information regarding the task solving processes children use (Resing, 2013). The main aim of the current thesis was to investigate the information that can be obtained from process-oriented dynamic testing and the specific characteristics of a new method of computer automated process assessment based on children's sequences of solving activities when answering cognitive problems.

Dynamic testing

Dynamic assessment or testing in its broadest definition constitutes an assessment procedure in which a form of feedback is provided to the child (Haywood & Lidz, 2007; Sternberg & Grigorenko, 2002). The various forms of feedback or training that are part of dynamic assessment or dynamic testing methods cover a whole range of instruction procedures (e.g., Haywood & Lidz, 2007; Sternberg & Grigorenko, 2002). Some methods, often referred to as "dynamic assessment" propose mediation that is provided on a highly individual basis (e.g., Feuerstein, Feuerstein, Falik, & Rand, 2002). Other dynamic assessment instruments, often referred to as "dynamic testing", were developed with a focus on standardized test-and-training formats (Sternberg & Grigorenko, 2002).

The graduated prompts method, utilized in this thesis, is a specific form of dynamic testing, which follows a pretest/training/posttest format,

aimed at finding the minimal number of prompts children need to solve tasks (Campione & Brown, 1978; Resing, 1993, 2000). The child is expected to work independently during the pre- and posttest, as help is not provided. During the training phase, graduated prompts approaches follow a standardized format, in which predetermined prompts are provided to children if and when they fail to answer a task correctly (Campione & Brown, 1978; Campione, Brown, Ferrara, Jones, & Steinberg, 1985; Resing, 1993, 2000). If a child is not able to solve a task correctly during training, at first, general, metacognitive prompts are provided. If a child cannot use these prompts to correctly solve the task, more specific, cognitive prompts will be provided. In a third phase, if children are still unable to solve the task correctly having been provided with these prompts already, a scaffolding procedure is offered, in which the process of solving the task is modelled to the child (Resing & Elliott, 2011; Resing, Xenidou-Dervou, Steijn, & Elliott, 2012). Both the number of prompts that children needed during training, and the number of items they could correctly solve at the posttest were found to be reasonably accurate predictors of children's learning potential (Caffrey, Fuchs, & Fuchs, 2008; Resing & Elliott, 2011).

The use of dynamic testing procedures has been shown in past research to provide additional explained variance regarding the prediction of school achievement outcomes when compared to the use of static tests only (Elliott et al., 2018; Grigorenko, 2008). However, computer automated process-oriented dynamic testing with the graduated prompts method, such as the tests used in this thesis, is a relatively new development, in relation to which much is still to be learned.

Process assessment

In psycho-educational research, several different methods for detecting and measuring task solving strategies and processes have been described (Ericsson, 2003; Tenison, Fincham, & Anderson, 2014), although these often represent a trade-off between the accuracy of the measurement, the reactivity it evokes from participants, and the ease of use for the tester (Tenison et al., 2014). In this context, reactivity concerns the influence of changes in the task solving situation on the task solving behavior of the child as a result of the measurement method (Kirk & Ashcraft, 2001; Tenison et al., 2014).

One method of process assessment is that of obtaining verbal reports (explanations) from participants, either during the solving process or immediately afterwards. This is generally seen as a valid method of obtaining data regarding the participants' task solving processes (Ericsson & Simon, 1980; Tenison et al., 2014). However, providing a verbal report during or after task solving might induce some change in the solving processes of participants (Kirk & Ashcraft, 2001; Tenison et al., 2014). Another method is that of time measurement. Recent literature, however, has indicated that the relationship between completion time and accuracy is not necessarily the same for all tasks, but instead might be influenced by the level of complexity of the task (Goldhammer et al., 2014; Scherer, Greiff, & Hautamäki, 2015).

Advances in computer technology, such as easier to use hardware, and more powerful software have created new possibilities for both recording data, and analyzing the actions captured in the data (Khandelwal & Mazalek, 2007; Verhaegh, Fontijn, Aarts, & Resing, 2013; Verhaegh, Hoonhout, & Fontijn, 2007). Despite the clear benefits of computerized testing, it also presents some pitfalls. Especially younger children are thought to experience difficulty using a PC mouse or keyboard interface (Verhaegh et al., 2013), including operations such as "drag-and-drop" procedures on touch-surface tablet interfaces (Price, Jewitt, & Crescenzi, 2015). Tangible User Interfaces (TUIs) provide a more natural modality for children, because these consist of tangible, physical materials, which are outfitted with sensors that enable detecting a child's manipulations of these materials (Verhaegh, Resing, Jacobs, & Fontijn, 2009).

In addition to the user interface, computerized systems need facilities to record data and translate the data into meaningful information or scores through some sort of interpretive step. Usually the data of a child's task solving behavior on a computerized task is saved into log-files (Goldhammer, Naumann, Rölke, Stelter, & Tóth, 2017). As Zoanetti and Griffin (2017) describe, the process of inferring meaningful information from all the actions recorded in the log-files has proven to be challenging. This thesis aimed to provide a rule-based theory-driven method of process assessment which is relatively easy to use. In some of our studies sensor technology was used to detect children's action sequences, which was thought to be an unobtrusive method for recording children's actions (Verhaegh et al., 2013).

Task solving processes

When a child is first presented with a task, he or she attempts to understand the task and how it has to be solved by creating an initial mental representation of the task (Fireman, 1996; Pretz, Naples, & Sternberg, 2003; Robertson, 2001). The efficiency and accuracy of the solving process is thought to be determined, in large part, by the quality of this representation (Robertson, 2001). Therefore, many researchers have argued that the initial representation is a crucial aspect of performance (Hunt, 1980; Pretz et al., 2003). It is often considered to be dependent on the availability of knowledge of the solver and the way this knowledge is organized. The representation changes with new information regarding the task becoming available (Pretz et al., 2003; Robertson, 2001).

A second factor that has been assumed to influence the task solving process is the familiarity of the solver with the type of task. The availability of prior models and experience with solving comparable tasks determines whether the solution can be attained through transfer of a previously used solving strategy, and whether or not the solver has global solving methods available that might lead to the solution (Newell & Simon, 1972; Weisberg, 2014). Newell and Simon (1972) stated that, if the solver has no clear strategy available, general methods such as “means-ends-analysis”, which consists of restructuring the representation through the division of the problem into sub-problems, can be used to find a solution (Robertson, 2001; Weisberg, 2015). Alternatively, strategies can be used to store information surrounding the task more efficiently in the mental representation. Halford and colleagues theorized that these strategies might consist of either combining information into meaningful units, referred to as conceptual chunking, or breaking up a task into simple steps that can be solved sequentially, referred to as the segmentation strategy (Andrews & Halford, 2002; Halford, Wilson, & Phillips, 1998). The task representation serves to organize the information surrounding the task, and is said to determine the strategies that are chosen to try and solve a task. Using inaccurate strategies may often be due to an incorrect representation. Moreover, learning to utilize new strategies has been argued to potentially improve a solver’s representation (Alibali, Phillips, & Fischer, 2009).

In his overlapping waves model of strategy use, Siegler (1996, 2007) described children's strategy use as quite variable, both within and between different trials. Even when the same task is presented repeatedly, children will often use different strategies to solve it. He posed that children choose a strategy that provides a balance between the fastest and most accurate solution. Similarly, accurate performance appears not to be related to any specific strategies, but instead to the ability to adaptively switch between strategies when the task demands it (Hunt, 1980). However, in the longer term, learning is characterized by increased use of more advanced strategies (Siegler & Svetina, 2002). Instability in strategy use is thought to indicate learning (Siegler, 2007).

Grouping of answer pieces (GAP)

A core subject of this thesis was the use of a new process-oriented measure, which was thought to provide information on children's task representations by analyzing how they grouped the different pieces in their sequence of answering. The idea of grouping of answer pieces was based on the basic principle of grouping information together into meaningful "chunks" to reduce the load of the task representation (Halford et al., 1998; Simon, 1974). In this thesis, this was executed using a rule-based algorithm to analyze the sequence of children's actions during the solving process. Pieces of the answer were considered grouped if they were completed in immediate succession of each other. In Appendix I more extensive explanation and operationalization was provided of the principles underlying this measure for the two tasks used in this thesis, an inductive reasoning task involving series completion and a complex figure task.

Inductive reasoning

The current thesis made use of inductive reasoning tasks, which require the solver to infer rules that govern a number of elements, described in terms of their similarities and differences (Klauer & Phye, 2008). A key component to the development of inductive reasoning ability and the level of complexity that children can handle in solving inductive reasoning problems seems to be the number of variables that can be encoded into a single mental

representation (De Koning & Hamers, 1999; Perret, 2015). The complexity of items might be reduced by the use of strategies that increase the number of variables that fit into a single representation (Halford et al., 1998; Halford, Wilson, & Phillips, 2010; Perret, 2015).

In several studies of this thesis, a series completion task was used that was developed and further adapted for computerized administration using TUIs (e.g., Resing & Elliott, 2011; Resing, Touw, Veerbeek, & Elliott, 2017). The task consisted of a figural series completion task, in which a series of puppets is presented, and the child is required to construct the last puppet figure using tangible blocks. To complete the tasks, the child has to identify how the patterns of the puppet pieces change throughout the series and determine the rule that governs these changes (Resing et al., 2012).

For series completion tasks, such as those used in this research, the process of solving series was described by, among others, Sternberg and Gardner (1983) and Sternberg (1985), who divided it into three phases. Initially, the solver has to encode the task into a mental representation, through perceiving the elements of the task and retrieving knowledge that is relevant to interpret them. Then, comparison takes place, which requires inferring the rule, mapping the higher-order rules that are involved, applying this rule and comparing the solution to the available options. Finally, the correct answer has to be justified. Alternatively, Simon and Kotovsky (1963) described the process as starting with scanning the series and attempting to detect the relationships governing the elements through formulating hypotheses. Once a relationship is detected, the length of a full cycle has to be discovered to establish the periodicity of the relationships. Then a rule can be formulated and, once the position of the answer is determined, applied to construct the final answer (Resing & Elliott, 2011).

The sequences of placement while constructing the answer (the correct puppet, consisting of eight separate pieces) were analyzed on grouping of answer pieces, through an automated algorithm that was based on task analysis of all the items in the test. Pieces that went through the same transformations, or shared common characteristics such as color, pattern, or anatomy, were considered grouped if they were placed immediately after each other in the sequence of solving (see Appendix I in the back of this dissertation for a more detailed account).

Complex figure drawing

This thesis also used a dynamic complex figure drawing task (Osterrieth, 1944; Rey, 1941), and reconstructed the grouping of answer pieces measure for use with this type of task. Complex figure drawing tasks can be used to assess visuo-spatial abilities, such as visual perception and construction (Martens, Hurks, & Jolles, 2014), and requires multiple cognitive processes, which include visuo-spatial perception and representation, planning, and working memory. Some researchers underlined the primary role of encoding for developmental changes in complex figure drawing ability, and pointed out that, in general, older children tend to solve the task through more configurational, coherent encoding of the figure (Kirkwood, Weiler, Bernstein, Forbes, & Waber, 2001). The algorithm for detecting grouping of answer pieces was adapted for use with the complex figure task. In this task, it was used to analyze which lines were drawn sequentially to use the configurational elements in structuring the representation of the task. Based on prior analyses of the processes involved in solving complex figure tasks (Kirkwood et al., 2001; Resch, Keulers, Martens, van Heugten, & Hurks, 2018), groups of lines were discerned that were considered useful when sequentially drawn together in time (i.e. drawn as one bigger configuration/cluster of lines, instead of as separate lines).

In this thesis, we explored whether children's grouping behavior would change as a result of training, and how grouping behavior in both the series completion and the complex figure task was related to task performance.

Outline of this thesis

The current thesis aimed to explore and examine aspects of process-oriented measurement in a dynamic testing context. Across all studies, both established process-oriented measures and a new measure were used, which was considered to measure children's grouping behavior in constructed response dynamic tests. The relations between task accuracy and task solving processes were investigated, as well as the effects of training on the sequences in task solving children employed to answer the tasks. The (additional) predictive value of the sequences in task solving for academic performance was topic of study as well.

Chapter 2 seeks to provide more insight into children’s task representations by analyzing their sequences of task solving steps. It was examined whether and how children, when solving series completion tasks, grouped the different task manipulatives into meaningful “chunks” (Halford et al., 1998). A tangible user interface (TUI) was used for studying natural task solving behavior by children in combination with computerized monitoring of the task solving steps. The study was executed with 8-year old children from regular school grade 2 classes, and children’s GAP scores, verbalizations, completion times, and proportion of time taken on the initial stages of the task were used to evaluate the relationship between the different measures and task accuracy. Additionally, the contributions of the task solving process measures to the prediction of item success were investigated.

Chapter 3 focuses on the effects of using a pretest in dynamic testing with a series completion task. In this study one group of 7-to 8-year old children received a dynamic test including a pretest and a second group of children a dynamic test without a pretest. The effects of utilizing a pretest on accuracy, GAP, verbalizations, and completion time were examined, including the use of a pretest on the number of prompts children need during a graduated prompts training procedure, and on the time children spend on the different stages of testing. Furthermore, this chapter explores the question whether the occurrence of a pretest leads to changes in the relationship between task solving processes in dynamic testing on one hand, and both task performance on the series completion task, and academic performance on math and reading comprehension on the other hand.

Chapter 4 presents 7-to-8-year old children’s progression in solving a dynamic test of series completion. The study employed a pretest-posttest control group design with randomized blocking. The effects of a graduated prompts training procedure on the accuracy scores on series completion, as well as shifts in the use of GAP, verbalized strategy use, and completion time were presented. In addition to the effects of training on processes children use while solving series completion tasks, this chapter also sheds a light on the predictive value of these task solving processes in relation to academic performance on math and reading comprehension.

Chapter 5 explores the use of sequential task solving steps (GAP) in a different domain, within a dynamic testing format. A dynamic test of complex figure drawing was administered to children in either a dynamic, graduated

prompts testing condition or a static unguided repeated testing condition. This chapter focuses on the effects and reach of a graduated prompts training on both performance and GAP measures of the dynamic complex figure task.

Chapter 6, the final chapter in this thesis, concludes with an overview of the results of the studies presented in the thesis, in order to synthesize the research findings into a first conclusion regarding the utility of process-oriented dynamic testing. Implications of these findings and recommendations for further research are highlighted.