



Universiteit Leiden

Psychologie  
Faculteit der Sociale Wetenschappen



# The purpose of the concept of Learning Potential

*The relation between Learning  
Potential and Intellectual Functioning  
in a primary school setting*

Renske de Haan

---

Master Thesis Child & Adolescent Psychology  
Date: November 2017  
Student number: S1592351  
Supervisor: J. Veerbeek MSc  
Second reader:

## Abstract

This study examined how learning potential is related to a child's intellectual functioning in a school setting. We executed a series completion task using a dynamic testing format, utilizing the graduated prompt approach. Learning potential was indicated by an electronic console which could measure completion time and accuracy, additionally learning potential was estimated by the teacher. Intellectual functioning in a school setting was evaluated by teachers. We hypothesized that learning potential would be related to intellectual functioning in a school setting, and that a part of learning potential could be estimated by the teacher. Participants were 176 children, ranging in age from 6 to 10 from primary schools in The Netherlands. We used a pre-test-post-test control-group block design. It was found that teachers could partly predict learning potential and that their prediction was strongly related to overall school performance and language performance. Above that, learning potential is somewhat related to school performance, yet it does not seem to be an adequate measure to sufficiently support the advice or prediction of school related aspects. Nevertheless, this purpose could be served by combining learning potential with other school related factors.

## List of contents

<b>Abstract</b>	<b>2</b>
<b>1. Introduction</b>	<b>4</b>
<b>2. Background</b>	<b>6</b>
<i>The concept of learning potential</i>	6
<i>Factors that correlate with learning potential</i>	7
<i>The graduated prompt technique</i>	7
<i>The zone of proximal development</i>	7
<i>Inductive reasoning</i>	8
<i>The electronic console</i>	9
<i>This research</i>	9
<b>3 Method</b>	<b>11</b>
3.1 <i>Participants</i>	11
3.2 <i>Design</i>	12
3.3 <i>Materials</i>	12
3.4 <i>Procedure</i>	15
3.5 <i>Scoring</i>	15
<b>4. Results</b>	<b>18</b>
4.1 <i>The effect of training</i>	18
4.2 <i>Teacher's estimation of learning potential</i>	20
4.3 <i>Learning potential in relation to school performance</i>	21
<b>5. Discussion</b>	<b>26</b>
5.1 <i>Effect of training</i>	26
5.2 <i>Teacher's estimation of learning potential</i>	27
5.3 <i>Learning potential in relation to intellectual functioning in a school setting</i>	28
5.4 <i>Overall conclusion</i>	30
5.5 <i>Limitations and recommendations</i>	31
<b>6. Literature</b>	<b>33</b>

## 1. Introduction

Knowledge transfer is one of the principle productive forces of our economic growth (Ordóñez & Sánchez, 2016). Consequently, investing in appropriate learning methods fitted to the child is a wise investment (Hartog, Oosterbeek & Teulings, 1993; Graczyk, Domitrovich, Small & Zins, 2006). Aiding children with learning difficulties and/or supporting the choice of an appropriate (middle)school could be aspects of investing in fitting learning methods. In this process, an estimation of a child's level of intellectual functioning can be helpful. This estimation can serve several purposes. As this estimation serves to support the choice of an appropriate (middle)school or to predict future school level, it is of great importance that this measure is related to a child's level of intellectual functioning in school setting. There are different procedures for estimating intellectual functioning of children.

Traditional, conventional or static testing is widely used and aims to provide an indication of a child's level of intellectual functioning (Bosma & Resing, 2012). These tests measure previously acquired knowledge, at a certain point in time, represented in an intelligence quotient (IQ; Kaldenbach, 2006). No feedback is given during a static test procedure. The IQ score has an average of 100 and a standard deviation of 15 and allows for comparison over time and between individuals. Therefore, IQ scores are widely used for educational placement and assessment of intellectual (dis)ability. The correct estimation of intellectual functioning and a child's school competence are very important as it determines their educational level and has major influence on their status later in life (McGrew & Wendling, 2010).

There are studies that suggest that the IQ score might not be sufficient in reflecting intellectual functioning. Specifically, IQ scores appear to be highly dependent on socioeconomic status (Bradley & Corwyn, 2002). Furthermore, IQ scores underestimate the cognitive abilities of children with learning disabilities (Siegel, 1989) or children from non-indigenous backgrounds (Nijenhuis, Willigers, Dragt & van der Flier, 2016). Above that, IQ scores do not provide information about the mistakes a child makes nor the effective way of helping the individual. Accordingly, the main applications of IQ tests in educational contexts are usually description, prediction and classification (Resing & Elliott, 2012). In contrast, dynamic testing proposes that *problem solving behaviour* and the *ability to learn* are more suitable measures for intellectual functioning (Haywood & Lidz, 2007). Therefore, it might

be more accurate to base the estimation of intellectual functioning, given its far-reaching consequences, on a child's problem solving behaviour and ability to learn.

Dynamic assessment procedures were developed as alternative measures to traditional intelligence tests. The principal characteristic of dynamic testing is based on the assumption that training during a test, including feedback and prompts, results in a more integrated indication of the level of intellectual functioning than static IQ tests (Resing, Touw, Veerbeek & Elliott, 2017). In research to dynamic testing, a pre-test – training – post-test format is often used (Resing, Touw et al., 2017). In the training-phase, the child is only assisted when it is not able to proceed independently. This assistance can be provided in diverse forms like prompts, hints or feedback originating from the principles of the graduated prompt technique. Within dynamic assessment, a child's level of intellectual functioning is given in terms of *learning potential*. This concept includes the progression as a result of the prompts, hints or feedback a child got. Children can show individual differences in progress when solving equivalent tasks. Some children are able to solve a task, imitating an example item. Other children need more examples and instruction to solve the same task. The amount of instruction, in combination with the child's progression could indicate a child's level of intellectual functioning (Lidz & Elliott, 2000; Resing, Touw et al., 2017). In relation to IQ tests, dynamic tests focus less on present cognitive functioning and more on the possible cognitive prospects of a child (Kolakowsky, 1998). Above that, learning potential fluctuates considerably between children, whereas differences between IQ scores between children are smaller (Bosma & Resing 2006). Therefore, IQ scores and learning potential needs to be interpreted independently.

Furthermore, dynamic testing gives insight in the way a child responds to several forms of feedback (Resing, Touw et al., 2017) and enables professionals to identify strengths and weaknesses in children's learning (Bosma & Resing, 2010). Given that the best way to help a child to learn, is to figure out the instructions to which the child is most responsive (Berk, 2001), dynamic testing may be valuable for helping both typically developing children and children with intellectual disabilities.

The approach to indicate learning potential, using dynamic test results, varies on multiple levels. For example, some studies combine pre-test scores with post-test scores in indicating learning potential, while others use post-test scores in isolation (Hessels, 2009). Also, the number of prompts needed in combination with post-test scores (Resing, Tunteler, De Jong,

& Bosma, 2009) and the number of prompts needed in isolation (Bosma & Resing, 2012) were used to identify learning potential.

Thus, methods of identifying learning potential fluctuate between studies. Consequently, it is complicated to predict how learning potential is related to intellectual functioning in school setting (Bosma & Resing, 2012). Research to this relationship has shown that dynamic test outcomes (e.g. the number of prompts a child needs to complete the task and their post-test accuracy scores) are accurate, individual predictors of future school success (Caffrey, Fuchs, & Fuchs, 2008). However, no research has been conducted to the relationship between learning potential and the child's current intellectual functioning in a school setting. Therefore, it is not clear whether learning potential can be used as a measure to support the advice or prediction of school related aspects (Bosma & Resing, 2012).

It is not yet known how learning potential is expressed in a child's daily school functioning. The current study aimed to investigate whether learning potential is an appropriate measure for the advice or prediction of school related aspects. The key objective of this study was to get insight in how learning potential is related to a child's intellectual functioning in a school setting.

## 2. Background

### The concept of learning potential

The scientific definition of learning potential seems to fluctuate. In her study to dynamic assessment, Kolakowsky (1998) describes learning potential as the ability to improve performance with practice. According to this definition, learning potential can only be measured within a multi-trial test procedure. Otherwise, improvement nor decline can be identified. The ability to benefit from instruction and the ability to generalize newly learned skills in a novel situation are not included in her definition. Other scientific definitions of learning potential are: the extent to which someone is able to benefit from instruction (Resing, Bakker, Pronk & Elliott, 2017) and the extent to which someone can accurately or strategically solve problems (Resing, Xenidou-Dervou, Steijn & Elliott, 2012).

## Factors that correlate with learning potential

Studies to the relation between learning potential and both a child's daily functioning and individual aspect of a child have been conducted. Equivalent to the current study, in these studies, learning potential is pointed out by a problem solving task. Learning potential appeared to be related to intelligence (Akbari & Hosseini, 2008), academic achievement (Greiff & Neubert, 2014) and strategy use (Resing & Elliott, 2011; Resing et al., 2012). However, the current research specifies in how learning potential is related to factors including a child's behaviour in daily life *in school setting*.

## The graduated prompt technique

In the graduated prompt technique, prompts are gradually provided to the child whenever it encounters problems in solving a task (Resing, 2000; Resing & Elliott, 2011). Following the concept of the graduated prompt technique, the first prompt is provided when a child is not able to succeed independently and gradually provide prompts until the child can solve the task (Resing, Touw et al., 2017). Hence, children are provided with the minimum amount of prompts necessary to progress on the task (Resing, Bakker, Pronk & Elliott, 2016). The sort of instruction a child needs to finish the task, indicates the kind of instruction the individual needs.

A lot of research has been conducted about dynamic testing following the graduated prompt technique. Training through administrating graduated prompts seems to result in greater accuracy, fewer corrections and reduced trial-and-error behaviour in a series completion task compared to repeated practise (Resing, Bakker et al., 2017; Resing, Touw et al., 2017; Resing et al., 2016). Dynamic training that follows the graduated prompt technique is related to more advanced problem solving behaviour (Resing & Elliott, 2011).

## The zone of proximal development

The graduated prompt technique is based on the concept of the zone of proximal development (ZPD; Vygotsky, 1980). The ZPD is an element of the sociocultural theory about the construction of knowledge from Vygotsky (Vygotsky, 1980). The ZPD refers to the difference between the level of performance a child is able to reach without guidance and the level of performance a child can reach when helped by someone with more understanding or skills in this field. Learning within the ZPD seems to be an effective way of learning (Vygotsky, 1980). For an adequate assessment of cognitive functioning, Vygotsky suggested

that an identification of both levels of performance is necessary (Kolakowsky, 1998). Thus, in order to learn effectively in this zone, children need help. The efficiency in which a child learns in the ZPD gives an indication for the learning potential, likewise the ability to benefit from instruction does (Kolakowsky, 1998).

### Inductive reasoning

Another important aspect in indicating cognitive ability is identifying the level of inductive reasoning (Goswami, 1996). Inductive reasoning concerns predicting (new) situations, based on earlier acquired knowledge. This procedure includes detecting a rule or a relation in a specific situation, generalize this rule, and subsequently apply it in several other (specific) situations (Raven, 2000). Research to cognitive abilities often use inductive reasoning tasks since inductive reasoning is necessary for learning and transfer (Ferrara, Brown & Campione, 1986). The cognitive processes in inductive reasoning consist of scrutinizing attributes of the objects or the relations between them; finding rules and regularities (Hayes, Heit & Swendsen, 2010; Resing, Touw et al., 2017). In everyday life, we make decisions and predictions based on this type of reasoning. People generalise knowledge from a specific situation to a more overarching situation. This generalization is a key component in learning about properties of an object, cause-effect relation, social rules and many other domains of knowledge, even learned at school (Tenenbaum, Griffiths & Kemp, 2006).

Klauer and Phye (2008) formulated a theory about inductive reasoning in which they made a distinction between two strategies of solving an inductive reasoning task: analytical and heuristic strategy. The superior, analytical strategy, consists of solving a task by planning, screening features and attributes and zooming in at the differences and similarities of the objects in order to find a rule. These skills seem to be accessible when solving an inductive reasoning task. On the other hand, heuristic strategy can be characterized by a global inspection of the task, followed by a quick solution that frequently appears to be based on trial-and-error. The more a child uses analytical strategy skills, the better it can solve novel problems by using rules based on prior problems (Crescentini, Seyed-Allaei, de Pisapia, Jovicich, Amati & Shallice, 2011). Helping children generalise by teaching them analytical strategies, ought to make them better at inductive reasoning. The feedback and instructions given in dynamic assessment should provide the guidance needed to solve inductive reasoning tasks.



### *Series completion task*

An example of an inductive reasoning task is a series completion task in which a logic sequence of objects should be finished. By seeking for similarities and differences between the objects, a rule that describes the changes between the objects can be found (Resing & Elliott, 2011). The inductive reasoning task used in this study is based on the task analytical model of series completion of Sternberg (1985).

### The electronic console

In this study a tangible user interface (TUI) has been used. The TUI contained an electronic board and digitally enhanced physical blocks which were detected by the board (Verhaegh, 2012). The TUI saved information about the child's performance and provided prompts as a reaction to this. Using a TUI enlarges the correctness of the scores. Above that, contrary to working with a mouse and keyboard, working with tangible objects provided more visual-spatial freedom (Olkun, 2003). The TUI is a concrete representation instead of a virtual representation of an object, resulting in the child showing more natural behaviour. However, using an electronic console may result in many data which are not directly interpretable, the translation into interpretable data is still a time-consuming process (Resing, Touw et al., 2017). Another disadvantage of using the TUI might be that for class wise administration, a school should purchase multiple TUI's, which could be costly. An alternative is to administer the task individually, though this is time consuming and requires an additional supervisor. However, investing in TUI's is a one-time investment.

### This research

In order to gain more insight in the effects of training on a child's problem solving behaviour and in how learning potential is related to a child's intellectual functioning in a school setting, this study aimed to provide an answer to the following points of concern. (1) The extent to which children will improve their problem solving behaviour, in response to training. (2) The extent to which teachers can predict learning potential. (3) The extent to which learning potential as (i) reported by dynamic testing and as (ii) estimated by the teacher is related to intellectual functioning in a school setting.

First, the way children improved their problem solving behaviour in response to training was investigated. It was expected that, due to training, children would become more accurate problem solvers and show less correcting behaviour during their response. This was

measured in terms of (a) accuracy, (b) completion time, (c) number of correct placed pieces and (d) number of corrections. A number of hypotheses was tested. (a) We expected that children in the trained condition would outperform the children in the control condition at the post-test accuracy scores (Resing & Elliott, 2011). (b) We expected that completion time would increase from pre-test to post-test for children in the experimental condition, but would stay unchanged for the children in the control condition. A comparable change in completion times was found for difficult items in a study with a similar design by Resing and Elliott (2011). The current study used the results of the difficult items since the series completion task in the study of Resing and Elliott (2011) seems easier; it consisted of less changing features than the one used in this study. (c) We expected that children who received dynamic training would place more pieces correctly than the children who did not receive dynamic training, measured at the post-test. Since earlier research found a diminishing number of corrections due to training (Resing, Touw et al., 2017), (d) we expected that the number of corrections would diminish for children in the trained condition from pre- to post-test, but not for children in the control condition.

Also, the relation between learning potential as reported by dynamic testing and learning potential estimated by the teacher, was examined. When the teacher's estimation of learning potential is equal to the dynamic test outcome, conducting a dynamic test procedure is not necessary anymore. We investigated the extent to which (a) the number of prompts needed, (b) pre- and (c) post-test accuracy scores were related to the learning potential estimated by the teacher. Out of these indicators for learning potential, a model that was best related to the teacher's estimation of learning potential was identified. Since there was no feedback nor training phase conducted within the pre-test, the pre-test accuracy scores were seen as a measurement of previous acquired knowledge and therefore evaluated as a static test score. Seeing that the number of prompts and the post-test accuracy scores are accurate individual predictors of future school success (Caffrey et al., 2008), we expected that out of the indicators of learning potential, a combination of the number of prompts needed to pass an item and the post-test accuracy scores (e.g. the dynamic test outcomes) would create a model that is best related to the teacher's estimation of learning potential.

Finally, we investigated how learning potential was expressed in a school setting. More precise; we examined how learning potential as (i) reported by dynamic testing and as (ii) estimated by the teacher was related to intellectual functioning in a school setting. Indicators for intellectual functioning in a school setting included: (a) overall school performance, (b)

mathematical performance, (c) language performance, (d) the need for instruction, (e) the static test score of mathematics and (f) the static test score of reading comprehension. Out of these indicators for intellectual functioning, a model that was best related to both learning potential as reported by dynamic testing and learning potential as estimated by the teacher, was identified. An assumption of research to dynamic testing is that registration of number and type of feedback and prompts during the test will result in a more integrated indication of learning potential (Resing, Touw et al., 2017). Therefore, we expected that out of the indicators for intellectual functioning in a school setting evaluated in this study, the need for instruction would be the greatest related indicator to both learning potential as reported by dynamic testing and learning potential as estimated by the teacher. Above that, since earlier research has found that learning potential is related to future school success (Caffrey et al., 2008), we expected to find the indicators including an impression of school performance (e.g. overall school performance, language performance and mathematical performance) related to learning potential as reported by dynamic testing. We expected that the indicators including an impression of school performance would also be related to the learning potential as estimated by the teacher, since language and mathematical tasks are daily practised tasks in a school setting. For this reason, the teacher is informed with the children's level of performance on these subjects. Consequently, the children's performance on these tasks might affect the teacher's evaluation of the child, including the overall school performance and the learning potential. Finally, we hypothesized that static test scores, to some extent developed for indicating school success (McGrew & Wendling, 2010), would also be related to both learning potential and teacher's estimation of learning potential.

## 3 Method

### 3.1 Participants

We recruited 176 children (90 boys and 86 girls) ranging in age from 6 to 10 ( $M = 7$  years 11 months,  $SD = 7$  months) from primary schools in towns in the western part of The Netherlands. All children attended regularly education classes. Parental consent for participation was obtained in all cases. Within the schools Dutch was the primary language.

**Table 1.** Design of the study

Condition	Raven	Pre-test	Training 1	Training 2	Post-test
Training	X	X	X	X	X
Control	X	X	-	-	X

*X: conducted*

*-: not conducted*

### 3.2 Design

This study utilized a previously studied format of dynamic testing: a pre-test-post-test control-group block design (table 1). To distribute children and their general cognitive ability randomly over both conditions, randomized blocking was performed. This blocking was based on Raven Standard Progressive Matrices scores which assessed the general cognitive ability of a child, explained in detail below (Conrad, 1976). Children were all seen four times individually. To children in the training condition a pre-test, two trainings and post-test was administered. To children in the control condition only a pre-test and post-test was administered, trainings were replaced by two dot-completion tasks. During the pre-test, a child was asked to solve a task without any assistance. Within the two trainings, help was granted in the form of instruction and feedback. In the post-test, as in the pre-test, no help was provided.

### 3.3 Materials

#### 3.3.1 Raven Standard Progressive Matrices

Randomized blocking was based on a measure of visual inductive reasoning: The Raven Standard Progressive Matrices (RSPM; Conrad, 1976; Raven, 2000). The task assessed the general cognitive ability of a child (Conrad, 1976). Consequently, the difference between the two conditions, could not be reproached to the children's general cognitive ability. An accuracy score from '0' – '60' was calculated for each child. The task consisted of 60 items divided over 5 sets that call upon children's ability to infer rules, the items within a set increased in difficulty. The RSPM comprised of black visual geometric designs on a white background, each with the same format: a 3 x 3 matrix in which the bottom right entry is missing. The child could choose from six to eight choices to fill in the missing piece. The task took approximately 20 minutes and was conducted class wise.

### 1.3.2 Dynamic test

#### - *Puppet series completion task*

The puppet series completion task was used as a measure for inductive reasoning. Children had to solve this visual-spatial series completion task, in which each item consisted of six puppets in line followed by a question mark. The items were presented in a paper booklet with one item per page. Each puppet was dressed a specific way, from which a pattern could be recognized. In addition, patterns in the puppet's gender could be perceived. The children had to construct the puppet that was supposed to be at the question mark. The puppet consisted of eight pieces: the head, two legs, two arms and three body parts. The head could be a boy's or a girl's head, directly describing the gender of the puppet. The rest of the pieces could be pink, yellow, green or blue. In addition, the pieces could be plain, dotted or striped. The pieces were represented as tangible blocks placed on an electronic console.

The degree of difficulty differed per item, depending on a several factors. (1) The number of changing features; a sequence of puppets in which the changing factor was the colour of the pants, should have been simpler to discover than a sequence of puppets in which the colours of the arms, legs, body parts and head differed between the puppets. (2) The period over which the sequence is repeated; an alternately repeated pattern should have been less complicated to find than a pattern that repeated every four puppets. (3) The last factor affecting the difficulty of an item was a combination of the two above. The hardest situation should have been one in which the puppets had many changing features that repeated within a dissimilar interval.

The pre- and post-test consisted of one example item and twelve test items. In the pre-test, the TUI first vocally explained that the child was supposed to finish the sequence (with an easy example item). The child was supposed to place the blocks on the electronic board. The TUI gave instruction when the child provided an inadequate answer to this example item. Thereafter, the child got a second trial. This was followed by twelve test items without any feedback or instruction. The pre- and post-tests started with the easiest item, slowly rising the degree of difficulty towards the last and hardest item. Due to this organization, the level of performance of a child without help was clearly shown by the amount of correctly solved items without help.

#### - *Training – graduated prompt technique*

For children in the trained condition, between the pre- and post-test, two trainings were administered. A training consisted of six test items. Contrary to the pre- and post-test

procedure, after an incorrect answer, feedback was provided and the child needed to try again. In total, a child got five trials to complete the item. The given feedback followed a structured scheme, according to the graduated prompt technique (Resing et al., 2017). The help consisted of a rising amount of instruction. After both the first and second incorrect answer, general metacognitive prompts were provided. The third and fourth incorrect answers were followed by a more task-specific cognitive prompt. The final prompt involved modelling of the solution process. Children did not get more than the individually needed feedback. Since the effect of every single prompt could be individually examined, this procedure provided specific information about the efficacy of the prompts. In the training sessions the items were slowly decreasing in difficulty to the last and easiest item. Nevertheless, in order to rehearse, the first item was the easiest. In this way, the child was able to apply its newly learned technique in a more accessible situation, potentially leading to fewer prompts. This structure provided insight in the extent to which the child could benefit from feedback: a component in learning potential (Resing, Bakker, Pronk & Elliott, 2017).

- *TUI*

To learn in the ZPD and to improve inductive reasoning, guidance had to be given. In order to compute a valid test procedure, every hint, prompt and reward was given in the exact same way to all children. This was achieved by using a TUI. The TUI is an A3-sized console developed by TagTiles to support independent learning (Verhaegh et al., 2017). The TUI contained an electronic table top sensing board with coloured light underneath it and audio output (Verhaegh, 2012). The child interacted with this board by using digitally enhanced physical plastic blocks which were detected by the board. This detection was facilitated by radio-frequency identification (RFID) tags in the tangible objects. Within the TUI and the objects, the visibility of the computer was reduced (Verhaegh, 2012). Children were supposed to place the blocks on the board to complete a puppet, after which the board supplied structured feedback. Dependent on the correctness of an answer, the TUI gave either a prompt or continued to the next item. Also, the TUI generated a variety of prompts: visual, verbal and (meta)cognitive. This exchange of information between the TUI and the child made it possible to gradually administer prompts. The TUI supported independent learning by children, since the amount of feedback was dependent on the performance of a child. In this way, the TUI executed the graduated prompt technique (Resing & Elliott, 2011) and provided detailed information on the problem solving processes (Verhaegh, Fontijn & Hoonhout, 2017; Resing & Elliott, 2011). The TUI saved the scores which were transferred

to a computer. The saved scores contained the position and correctness of the blocks, the timing of the responses and the amount of corrections.

### *3.3.3 Teacher's questionnaire*

Teachers were asked to fill in one questionnaire for each child concerning the impression of the child's school performance, static test scores (Cito-scores) and their impression about the child's need for instruction. The questionnaire consisted of 10 items about school performance and 13 items about their need for help. Teachers were asked to evaluate the child by comparing it to its peers, not only in its own class but with all the peers the teacher knew. Filling in one questionnaire took approximately 10 minutes per child.

## 3.4 Procedure

The first test administered was the Raven Standard Progressive Matrices, which was class wise administered in each classroom. Based on their scores on this task, children were blocked into the experimental- or control condition. The average cognitive ability of the children was equal in both conditions. Thereafter, the dynamic test procedure was started. Each child was individually tested for four times: pre-test, either two training sessions or two control sessions and post-test. These sessions took place in a separate room in the school. The TUI led the program and gave the prompts. Meanwhile the mentor was scoring the answers to make sure no data could be lost in case the computer crashed. The mentor escorted the child from and to class and made sure the child was paying attention to the sessions. One session took approximately 25-40 minutes and were given in intervals of 3-10 days.

## 3.5 Scoring

The data collected by the electronic console was transferred to a computer, recoded into numeric data and then converted into SPSS for analysis. Features of interest for each item and child were: accuracy, number of prompts needed, completion time, correct placed number of pieces and number of corrections.

### *3.5.1 Accuracy*

Accuracy was measured for each item. Accuracy included the number of items solved completely correct. Thus, when a child had the clothing of the puppet correct, but the gender incorrect, the item was scored as incorrect. The scoring of accuracy was binary; a correct

answer was coded as '1', an incorrect answer was coded as '0'. This system made it easy to picture the total correct answered items by summing up the scores. The accuracy scores were calculated with a number between '0' - '12' as the pre- and post-test consisted both of twelve test items.

### *3.5.2 Number of prompts*

The number of prompts was only measured in the training sessions. The amount of prompts was calculated for each item. This was expressed in a number between '0' - '4': '0' indicated that no prompts were needed to pass this item, '4' indicated that four prompts were needed to pass this item etc. By summing up these scores, the total number of prompts during two training sessions per child was determined varying from '0' - '24' (6 items x 4 prompts).

### *3.5.3 Completion time*

The time the child needed to complete one item is called the completion time. This time was measured from the moment the child saw the sequence of puppets, until the moment the final block was laid down. In the pre- and post-test this resulted in twelve completion times each. Summing up these times over pre- and post-test, resulted in a total completion time for each child. Completion time was calculated in milliseconds.

### *3.5.4 Number of correct placed pieces*

The number of correct placed pieces was measured in terms of the total accurately placed pieces over all the placed pieces. The total number of placed pieces was 96. Therefore, the number of correct placed pieces varied from 0 - 96.

### *3.5.5 Number of corrections*

The number of corrections referred to the number of times a child changed his or her given answer in order to improve it. No distinction was made between correct and incorrect corrections.

### *3.5.6 Teacher's estimation of learning potential*

The child's learning potential estimated by the teacher was obtained from the teacher's questionnaire. Teachers had to evaluate the child on a 5-point scale from high to low. Every child was assigned to a level indicating their learning potential, illustrated with a digit from



'1' to '5' following the teacher's estimation. Children who scored within the best performing 20% scored a '5'; 21-40% a '4'; 41-60% a '3'; 61-80% a '2'; and children scoring in the lowest 20% scored a '1'. High scores on this scale correspondent with high estimation of learning potential.

### *3.5.7 School performance*

The overall school performance, mathematical performance and language performance were also obtained from the teacher's questionnaire. This consisted of impressions by the teacher. These impressions were evaluated on the same 5-point scale as the learning potential estimated by the teachers. High scores on this scale correspondent with high estimation of school performance.

### *3.5.8 Need for instruction*

An estimation of the children's need for instruction was obtained from the 13 items teacher questionnaire. Each item consisted of two statements, referring to oppositional behaviour. First, the teacher needed to decide which statement applied to the child the best. Second, the degree to which this statement seemed an appropriate reflection of the child's behaviour had to be determined on a 3-point scale. Therefore, each item had six potential outcomes, scored on a scale from '-3' to '3' ('-3', '-2', '-1', '1', '2' and '3'). Zero was not a possible outcome, since the teacher had to choose between the statements. For each child, these 13 outcomes were summed up, which indicated their level of need for instruction given in a digit between '-39' and '39' (thirteen questions, each with six possible answers). Within six of the thirteen items, negative values presented a low need for instruction, therefore, these six had to be recoded ('-3' ↔ '3'; '-2' ↔ '2'; '-1' ↔ '1'). High scores on this scale correspondent with a low need for instruction and the capability of working independently.

### *3.5.9 Cito-scores*

The Cito-scores refer to scores on a commonly used static test in the Netherlands: Cito test (Hoijtink, Béland & Vermeulen, 2014). Cito tests identify the level of performance on several school subjects (Hollenberg, Van der Lubbe & Sanders, 2017). There are different Cito tests for each of the school subjects. Generally, these tests are administered three times a year (at the beginning, middle and end of the study year). The most recent scores available were included in this research. The Cito tests were not administered in the current study, the

outcomes were asked for in the teacher's questionnaire. Cito-scores used in this study were those of arithmetic and reading comprehension.

The Cito-scores were classified in 5 categories from 'A' to 'E'. Children who scored within the best performing 25% (75-100%) of the Netherlands got an 'A'; 50-74% a 'B'; 25-49% a 'C'; 11-24% a 'D' and 0-10% an 'E' (Jolink, Tomesen, Hilde, Weekers & Engelen, 2015; Janssen, Hop & Wouda, 2015). In order to compare Cito-scores with other variables in this study, the categories from 'A' to 'E' were translated to digits from '1' to '5': 'A' → '5'; 'B' → '4'; 'C' → '3'; 'D' → '2' and 'E' → '1'. Thus, for each child, static test performance was illustrated with a digit from '1' to '5'.

## 4. Results

Before analysing the data, the differences between children in the dynamically trained condition and children in the control condition were considered. Results of two one-way ANOVA's revealed no differences in the initial level of inductive behaviour (based on the scores of the Raven Standard Progressive Matrices (Raven, 2000,  $F(1, 174) = 1.57$ ;  $p = .212$ ) nor age ( $F(1, 174) = .053$ ;  $p = .817$ ) between the conditions. The control condition existed of 87 children; 43 girls and 44 boys. The dynamically trained condition existed of 89 children; 43 girls and 46 boys.

### 4.1 The effect of training

First, we investigated whether children could, in response to training, improve their series completion problem solving behaviour. Performances on the pre- and post-test were compared between children in the trained condition and children in the control condition. This was measured in terms of (a) accuracy, (b) completion time, (c) number of correct placed pieces and (d) number of corrections. Descriptive statistics are presented in table 2 (mean, SD at pre-test and post-test for the trained and untrained condition). Several repeated measures ANOVA's were run with session (pre-test/post-test) as within-subject variable and condition (control/training) as between-subject factor.

**Table 2.** Mean scores (*M*) and standard deviations per condition and session.

	N	Control condition		Trained condition	
		Pre <i>M</i> (SD)	Post <i>M</i> (SD)	Pre <i>M</i> (SD)	Post <i>M</i> (SD)
Accuracy	176	4.82 (2.22)	5.57 (2.75)	5.20 (2.39)	7.33 (2.54)
Completion time	146	875347.19 (220176.11)	824149.29 (215775.02)	894797.64 (223335.87)	818208.30 (233796.90)
No. of pieces correct	175	76.10 (11.73)	77.59 (11.80)	78.06 (10.89)	83.51 (10.10)
No. of corrections	173	1.97 (2.17)	1.49 (2.06)	1.74 (2.25)	1.51 (1.93)

*(a) Accuracy*

We expected that children in the trained condition would show greater progression from pre- to post-test in accuracy scores than untrained children. For accuracy scores, the repeated measures ANOVA showed a significant within factor main effect for session ( $F(1, 174) = 63.11; p < .001; \eta^2 = .27$ ). It did show a significant between factor main effect for condition ( $F(1, 174) = 10.67; p = .000; \eta^2 = .06$ ). More importantly, the analysis showed a significant within subject interaction effect for session x condition ( $F(1, 174) = 14.16; p < .001; \eta^2 = .08$ ). As expected, dynamically trained children showed a greater progression in accuracy scores from pre- to post-test than the untrained children.

*(b) Completion time*

We expected that children's completion time would increase due to dynamic training. For completion times, results of the repeated measures ANOVA showed a significant within factor main effect for session ( $F(1, 144) = 15.42; p < .001; \eta^2 = .097$ ). However, seeing the diminished completion times in table 4, this indicates that children in both conditions reduce their completion time. No between factor main effect for condition ( $p = .839$ ), nor within factor interaction effect for session x condition is shown ( $p = .437$ ) by the repeated measures ANOVA. Therefore, we can conclude that the reduction of completion time did not differ significantly between the conditions.

*(c) Correct placed pieces*

We expected that children who received dynamic training would make more progression from pre- to post-test at amount correctly placed body parts than the children who did not receive dynamic training. More precise: we expected that children in the trained condition would score a greater amount of correctly placed body parts than the children in the control

condition, measured at the post-test. For the amount of correct placed pieces, results of the repeated measures ANOVA showed a significant within factor main effect for session ( $F(1, 173) = 25.51; p = .000; \eta^2 = .129$ ) and a significant between factor main effect for condition ( $F(1, 173) = 6.54; p = .011; \eta^2 = .036$ ). Additionally, a significant within factor interaction effect for session x condition was found for the correct placed pieces ( $F(1, 173) = 8.32; p = .004; \eta^2 = .046$ ). We can conclude that children in both conditions made progression in the amount of correctly placed body parts from pre-test to post-test. Above that, for dynamically trained children this progression was greater than for the untrained children.

*(d) Number of corrections*

We expected that the number of corrections would diminish for the children in the trained condition, but not for the children in the control condition. Contrary to our expectation, for number of corrections, the repeated measures ANOVA showed no significant main effect for session ( $p = .071$ ), nor for condition ( $p = .682$ ). Finally, no significant interaction effect for session x condition was found ( $p = .526$ ). We can conclude that both dynamically trained children and untrained children did not change their number of corrections from pre- to post-test.

## 4.2 Teacher's estimation of learning potential

We investigated the extent to which (a) the number of prompts needed to pass an item, (b) pre- and (c) post-test accuracy scores were related to the learning potential estimated by the teacher. We expected that out of these indicators of learning potential, a combination of the number of prompts needed to pass an item and the post-test accuracy scores would create a model that is best related to the teacher's estimation of learning potential.

**Table 3.** Regression analysis (dependent: the learning potential estimated by the teacher).

Model	Adjusted R Square	R Square Change	Sig. F Change
1	.225	.235	.000
2	.264	.048	.028

Predictors: (constant), accuracy at post-test

Predictors: (constant), accuracy at post-test, accuracy at pre-test

A stepwise linear regression analysis, in which the dependent variable was the learning potential estimated by the teacher and predictors were accuracy at pre-test, accuracy at post-test and total number of prompts needed in the training sessions, was conducted. Table 3 displays the results. This stepwise analysis revealed two significant predictive models. The first model included solely accuracy at post-test as a predictor, which explained 22.5% of the variances of the learning potential estimated by the teacher ( $R^2 = .235$ ;  $F(1, 76) = 23.39$ ;  $p = .000$ ). The contribution of accuracy at post-test, holding all other variables constant, was positive: the more tasks solved accurately at the post-test, the higher the learning potential estimated by the teacher ( $Beta = .485$ ;  $p = .000$ ).

By adding pre-test accuracy as an additional predictor, the second model was created, by which 26.4% of the variances of the learning potential estimated by the teacher was explained ( $R^2 = .284$ ;  $F(1, 75) = 14.84$ ;  $p < .05$ ). The contribution of these indicators, holding all other variables constant was again positive. The more tasks solved accurately at the post-test ( $Beta = .365$ ;  $p = .002$ ) and at the pre-test ( $Beta = .250$ ;  $p = .028$ ), the higher the learning potential estimated by the teacher

When adding the total number of prompts needed in the training sessions, the predicted explanatory value of variances of the learning potential estimated by the teacher does not increase. Therefore, it can be concluded that, contrary to our expectations, a combination of accuracy scores on pre- and post-test created a model that is best related to the teacher's estimation of learning potential. This model showed that the combination of a static and a dynamic test outcome were best related to the learning potential estimated by the teacher.

### 4.3 Learning potential in relation to school performance

Finally, intellectual functioning in a school setting was examined. The extent to which learning potential as reported by (i) dynamic testing and as (ii) estimated by the teacher was related to intellectual functioning in a school setting was examined. Indicators for intellectual functioning in a school setting were: (a) overall school performance, (b) mathematical performance, (c) language performance, (d) the need for instruction, (e) the static test score of mathematics and (f) the static test score of reading comprehension. We expected that all indicators for intellectual functioning would be related to both learning potential and the teacher's estimation of learning potential. We expected the indicator need for instruction, would be the one best related to both indications of learning potential. Due to missing data

spread across the variable, the analysis is split into three independent regression analyses: (1) considerations of the teacher, (2) need for instruction and (3) static test scores.

*(i) Learning potential as reported by dynamic testing in relation to intellectual functioning in a school setting*

Learning potential as reported by dynamic testing is operationalized as post-test accuracy scores. Several stepwise linear regression analyses were performed. The dependent variable was post-test accuracy scores. Predictors were the indicators for intellectual functioning in a school setting, as mentioned above. Results of the stepwise regression analyses are shown in table 4.

The first regression analysis (4.1) concerned the considerations of the teacher. Included in this analysis were the indicators (a) overall school performance, (b) mathematical performance and (c) language performance. This analysis showed two predicting models. The first model included the indicator mathematical performance and explained 9.8 % of the variances of learning potential as reported by dynamic testing ( $R^2 = .103$ ;  $F(1,167) = 19.18$ ;  $p = .000$ ). Within this model, mathematical performance was positively related to learning potential, holding all other indicators constant ( $Beta = .321$ ;  $p = .000$ ). The higher the mathematical performance, the higher the learning potential reported by dynamic testing. The second model included the indicators mathematical performance and language performance and explained 11.6 % of the variances of learning potential as reported by dynamic testing ( $R^2 = .126$ ;  $F(2,166) = 12.00$ ;  $p = .000$ ). Within this model, both mathematical performance ( $Beta = .212$ ;  $p = .000$ ) and language performance ( $Beta = .187$ ;  $p = .000$ ) were positively related to learning potential, holding all other indicators constant. The higher the mathematical and language performance, the higher the learning potential reported by dynamic testing.

The second regression analysis (4.2) included the indicator (d) the need for instruction. This analysis showed a significant model. The model included the indicator need for instruction and explained 5.2% of the variances of learning potential as reported by dynamic testing ( $R^2 = .059$ ;  $F(1,145) = 9.02$ ;  $p = .003$ ). Within this model, need for instruction was positively related to learning potential, holding all other indicators constant ( $Beta = .242$ ;  $p = .003$ ). The

higher the scores on the indicator need for instruction (implies the capability of working independently), the higher the learning potential reported by dynamic testing.

**Table 4.** Regression analysis (dependent: the learning potential as reported by dynamic testing)

<i>4.1 Included in this model: (a) overall school performance, (b) mathematical performance, (c) language performance</i>			
Model	Adjusted R Square	R Square Change	Sig. F Change
1	.098	.103	.000
2	.116	.023	.037
Predictor model 1: (constant), mathematical performance Predictors model 2: (constant), mathematical performance, language performance			
<i>4.2 Included in this model: (d) the need for instruction</i>			
Model	Adjusted R Square	R Square Change	Sig. F Change
1	.052	.059	.003
Predictor model 1: (constant), the need for instruction			
<i>4.3 Included in this model: (e) the static test score of mathematics, (f) the static test score of reading comprehension</i>			
Model	Adjusted R Square	R Square Change	Sig. F Change
1	.064	.073	.004
Predictor model 1: (constant), the static test score of mathematics			

The third regression analysis (4.3) concerned the static test scores. Included in this analysis were the indicators (e) the static test score of mathematics and (f) the static test score of reading comprehension. This analysis showed one significant model. The model included the indicator the static test score of mathematics and explained 6.4% of the variances of learning potential as reported by dynamic testing ( $R^2 = .073$ ;  $F(1,112) = 8.76$ ;  $p = .004$ ). Within this model, the static test score of mathematics was positively related to learning potential, holding all other indicators constant ( $Beta = .269$ ;  $p = .004$ ). The higher the static test score of mathematics, the higher the learning potential reported by dynamic testing.

*(ii) Learning potential estimated by the teacher in relation to intellectual functioning in a school setting*

A stepwise linear regression analysis was used to investigate what indicators for intellectual functioning in a school setting correlate with the teacher's estimation of learning potential. The dependent variable was the teacher's estimation of learning potential. Predictors were the indicators for intellectual functioning in a school setting, as mentioned above. This stepwise

analysis revealed three significant predictive models. Results of this stepwise regression analysis are shown in table 5.

**Table 5.** Regression analysis (dependent: the learning potential estimated by the teacher)

<i>5.1 Included in this model: (a) overall school performance, (b) mathematical performance, (c) language performance</i>			
Model	Adjusted R Square	R Square Change	Sig. F Change
1	.711	.712	.000
2	.746	.037	.000
Predictor model 1: (constant), overall school performance Predictors model 2: (constant), overall school performance, language performance			
<i>5.2 Included in this model: (d) the need for instruction</i>			
Model	Adjusted R Square	R Square Change	Sig. F Change
1	.141	.148	.000
Predictor model 1: (constant), the need for instruction			
<i>5.3 Included in this model: (e) the static test score of mathematics, (f) the static test score of reading comprehension</i>			
Model	Adjusted R Square	R Square Change	Sig. F Change
1	.334	.340	.000
2	.420	.091	.000
Predictor model 1: (constant), static test score of mathematics Predictors model 2: (constant), static test score of mathematics, static test score of reading comprehension			

The first regression analysis (5.1) concerned the considerations of the teacher. Included in this analysis were the indicators (a) overall school performance, (b) mathematical performance and (c) language performance. This analysis showed two significant models. The first model included the indicator overall school performance and explained 71.1 % of the variances of learning potential estimated by the teacher ( $R^2 = .712$ ;  $F(1,154) = 381.57$ ;  $p = .000$ ). Within this model, overall school performance was positively related to learning potential, holding all other indicators constant ( $Beta = .844$ ;  $p = .000$ ). The higher the overall school performance, the higher the learning potential estimated by the teacher. The second model included the indicator overall school performance and language performance and explained 74.6 % of the variances of learning potential estimated by the teacher ( $R^2 = .750$ ;  $F(2,153) = 229.12$ ;  $p = .000$ ). Within this model, overall school performance ( $Beta = .606$ ;  $p = .000$ ) and language performance ( $Beta = .307$ ;  $p = .000$ ) were positively related to learning potential estimated by the teacher, holding all other indicators constant. The higher the



overall school performance and language performance, the higher the learning potential estimated by the teacher.

The second regression analysis (5.2) included the indicator (d) the need for instruction. This analysis showed a significant model. The model included the indicator need for instruction and explained 14.1 % of the variances of learning potential estimated by the teacher ( $R^2 = .148$ ;  $F(1,131) = 22.74$ ;  $p = .000$ ). Within this model, the need for instruction was positively related to learning potential estimated by the teacher, holding all other indicators constant ( $Beta = .385$ ;  $p = .000$ ). The higher the scores on the indicator need for instruction (implies the capability of working independently), the higher the learning potential estimated by the teacher.

The third regression analysis (5.3) concerned the static test scores. Included in this analysis were the indicators (e) the static test score of mathematics and (f) the static test score of reading comprehension. This analysis showed two significant models. The first model included the indicator static test score of mathematics and explained 33.4% of the variances of learning potential estimated by the teacher ( $R^2 = .340$ ;  $F(1,112) = 57.63$ ;  $p = .000$ ). Within this model, the static test score of mathematics was positively related to learning potential, holding all other indicators constant ( $Beta = .583$ ;  $p = .000$ ). The higher the static test score of mathematics, the higher the learning potential estimated by the teacher. The second model included the indicator static test score of mathematics and static test score of reading comprehension and explained 42.0 % of the variances of learning potential estimated by the teacher ( $R^2 = .431$ ;  $F(2,111) = 41.96$ ;  $p = .000$ ). Within this model, static test score of mathematics ( $Beta = .403$ ;  $p = .000$ ) and reading comprehension ( $Beta = .351$ ;  $p = .000$ ) were positively related to learning potential, holding all other indicators constant. The higher the static test score of mathematics and reading comprehension, the higher the learning potential estimated by the teacher.

For further exploration of the results, a fourth regression analysis in which the dependent variable was teacher's estimation of learning potential was performed. This analysis included the indicators a, b, c, and d. In this analysis, the additive value of (d) the need for instruction to the first analysis including the considerations of the teacher (4.1) was investigated. Results are shown in table 6.

**Table 6.** Regression analysis (dependent: the learning potential estimated by the teacher)

5.4: Included in this model: (a) overall school performance, (b) mathematical performance, (c) language performance, (d) need for instruction

Model	Adjusted R Square	R Square Change	Sig. F Change
1	.708	.711	.000
2	.735	.029	.000

Predictors model 1: (constant), overall school performance

Predictors model 2: (constant), overall school performance, need for instruction

This fourth regression analysis (5.4) included the indicators (a) overall school performance, (b) mathematical performance, (c) language performance and (d) need for instruction. This analysis showed two significant models. The first model included the indicator overall school performance and explained 70.8 % of the variances of learning potential estimated by the teacher ( $R^2 = .711$ ;  $F(1,125) = 306.96$ ;  $p = .000$ ). Within this model, overall school performance was positively related to learning potential estimated by the teacher, holding all other indicators constant ( $Beta = .843$ ;  $p = .000$ ). The higher the mathematical performance, the higher the learning potential estimated by the teacher. The second model included the indicator mathematical performance and need for instruction explained 73.5 % of the variances of learning potential estimated by the teacher ( $R^2 = .739$ ;  $F(2,124) = 175.78$ ;  $p = .000$ ). Within this model, overall school performance ( $Beta = .635$ ;  $p = .000$ ) and language performance ( $Beta = .268$ ;  $p = .000$ ) were positively related to learning potential estimated by the teacher, holding all other indicators constant. The higher the overall school performance and language performance, the higher the learning potential estimated by the teacher.

## 5. Discussion

It is complicated to predict how learning potential is related to intellectual functioning in a school setting. This relation is crucial, when learning potential serves as an indicator for advising or predicting school related aspects. Accordingly, it is essential to know how learning potential is related to school related aspects.

### 5.1 Effect of training

Several authors have demonstrated that children can improve their problem solving behaviour following dynamic training (Lidz & Elliott, 2000; Bosma & Resing, 2012; Resing, Touw et al., 2017). The outcomes of the present study support these findings. Dynamic training

resulted in both greater accuracy and higher proportions of correct placed pieces. However, our dynamic training procedure did not result in increasing completion times nor in reduction of corrections. Contrary to our expectation, the completion time did not increase correspondingly to the completion time in the difficult items in the study of Resing and Elliott (2011). Alternatively, for children in both conditions the completion time reduced, even so did the overall completion time in the study of Resing and Elliott (2011). Apparently, the series completion task used in this study is comparable to the one in the study of Resing and Elliott (2011). An explanation for the decreasing, instead of the expected increasing completion times, could be that practising by means of repetition has sped up the reaction process resulting in decreased completion times (Light, Reilly, Behrman & Spirduso, 1996). Notably, earlier research found a diminishing number of corrections in both conditions, which differed between both conditions (Lidz & Elliott, 2000; Bosma & Resing, 2012; Resing, Touw et al., 2017). The present study did not find a reduction in correction behaviour in conditions, nor between conditions. It could be that within our dynamic testing procedure, children may have felt the pressure to achieve, more than in earlier research, and therefore continued correcting their first given answers, more than in earlier research. Factors that could have contributed to the pressure to achieve might be time oriented or strict expressing researchers or idealistic school expectations.

In line with earlier research (Lidz & Elliott, 2000; Bosma & Resing, 2012; Resing, Touw et al., 2017), we can conclude that our dynamic test procedure was indicating learning potential, since the procedure could distinguish children between the two conditions based on their test scores. Above that, we can conclude that the graduated prompt technique in dynamic testing resulted in improvement of children's problem solving behaviour.

## 5.2 Teacher's estimation of learning potential

We examined the extent to which dynamic test scores, used to identify learning potential, were related to the teacher's estimation of learning potential. Based on earlier research (Caffrey et al., 2008), we expected that out of the indicators of learning potential, a combination of the number of prompts needed to pass an item and the post-test accuracy scores would create a model that is best related to the teacher's estimation of learning potential. Contrary to our expectations, instead of the number of prompts needed, the pre-test accuracy scores in combination with post-test accuracy scores created a model that was best related to the teacher's estimation of learning potential. Teacher's estimation of learning

potential was related to test outcomes which both represent a level of performance, rather than an amount of help a child needed to achieve this level. An explanation for this could be that pre- and post-test accuracy scores represent levels that are quantified and (therefore) unambiguous. The amount of help, on the other hand, is not quantified by school or a test (as this is an estimation of the teacher). A possible result of this is that teachers have an inadequate concept of a child's need for instruction.

We can conclude that the teacher is partly able to predict learning potential. Since a combination of dynamic and static test scores created a model that was best related to the teacher's estimation of learning potential, this study confirms the value of dynamic test scores supplementary to static test scores and the teacher's estimation.

### 5.3 Learning potential in relation to intellectual functioning in a school setting

To shed light on how learning potential is expressed in a school setting, we examined how learning potential as (i) reported by dynamic testing and as (ii) estimated by the teacher is related to intellectual functioning in a school setting. The results of this study showed that, contrary to our expectations, the child's need for instruction is the slightest related to learning potential. This finding interferes with the assumption of dynamic testing that the resources a child needs to achieve a certain level is a valuable indicator for learning potential (Resing, Touw et al., 2017). A possible explanation for the finding that the amount of help is the least related to learning potential is that the need for instruction could not have been a valid operationalization for children with an actual high need for instruction. While help seeking may arise from several motives. A motive could be that the child prefers someone else to solve the problem (Nelson-Le Gall & Glor-Scheib, 1985). Help seeking is related to persistent overall self-esteem and mastery goal orientation (Karabenick & Kanpp, 1991; Gonida, Karabenick, Makara & Hatzikyriakou, 2014; Carr, Luckin, Yuill & Avramides, 2013). Above that, high achieving students seek for help more frequently than low achieving students (Lui, 2009). It may also be that teachers did not acknowledge children with higher need for instruction, or labeled them as less intelligent.

Furthermore, our test results showed that learning potential as reported in dynamic testing is limitedly related to the static test scores of mathematics and not related to the static test scores of reading comprehension. Teacher's estimation of learning potential is related to static test scores of both mathematics and reading comprehension.

Following our test results, the consideration of the teachers, including indicators (a) overall school performance, (b) mathematical performance and (c) language performance, was best related to learning potential as reported by dynamic testing and as estimated by the teacher. This is in line with earlier research that has found that learning potential is related to future school success (Caffrey et al., 2008). However, the contribution of the indicators varied between learning potential as reported in dynamic testing and learning potential as estimated by the teacher. Learning potential as reported by dynamic testing is limitedly related to mathematical performance and even less to language performance. The relation to mathematical performance (and mathematical static test scores) can be partly comprehended by the mathematic nature of the task: seeking patterns is ranked to the mathematical thinking (Sfard, 1991). Above that, hardly no reading or language knowledge is required for passing the items.

It seems remarkable that the impression of overall school performance was not related to learning potential as reported by dynamic testing, whereas both language and mathematical performance were. Learning potential was related to school performances for the subject's language and mathematics, but possibly less to performances on the remaining school subjects, such as history, topography, traffic, geography and biology. It is possible that an overarching factor is affecting both the performances on language and mathematics and simultaneously affecting the performance on the series completion task used in this study. This overarching factor could be the use of an analytical strategy. As found in earlier studies, analytical strategy use contributes to a high reported learning potential by dynamic testing (Resing & Elliott, 2011; Resing et al., 2012). This analytical strategy use could be also beneficial within language and mathematics since using cognitive behaviour strategy is related to math achievement, mathematical problem solving, writing and reading performances (Eshel & Kohavi, 2003; Özsoy & Ataman, 2009; Lu & Liu, 2015; Shawer, 2016). Though, this strategy use might not be as such beneficial within the remaining school subjects.

Nearly three quarters of the variance of learning potential as estimated by the teacher can be explained by overall school performance and language performance. The teacher's estimation of learning potential was considerably related to overall school performance and limitedly to language performance. Teachers hardly reported children with high learning potential and low school performance (underperforming children) nor children with a low learning potential and a high school performance. According to the teachers, children

perform at a level which could be expected from them, based on their learning potential. Presumably, teachers are not able to distinguish between learning potential and level of performance. Above that, the theory that the estimation of the performance on a language and mathematical task would influence the estimation of the overall school performance can be rejected. Adding the need for instruction barely extends these three quarters. Possibly, teachers evaluate lower performing children as children with a higher need of instruction, which could result in the event that the indicator need for instruction is covered by the indicator overall school performance. It is remarkable that the teacher's estimation of language performance was related to the teacher's estimation of learning potential, while this was not an area to which learning potential as reported by dynamic testing was related. A reason for this inconsistency can be that teachers could have evaluated children with high language performance right away as children with high learning potential, as verbally strong children are often overestimated by their environment (Kaldenbach, 2006).

The learning potential estimated by the teacher was more related to intellectual functioning in a school setting than learning potential as reported by dynamic testing. This is not surprising seeing that intellectual functioning in a school setting is estimated by the teacher, as is learning potential. The total impression of a child and personal preference of the teacher could have similarly influenced both factors.

According to our test results, a high potential to learn is not equivalent to high performance in school. Nevertheless, more factors could affect school performance. These factors might cover (a part of) the relation between learning potential and intellectual functioning in a school setting. Race, socio-economic status, the community's social stock, motivation and attention are factors that influence primary school performance (Misra, Grimes & Rogers, 2013; Corpus & Wormington, 2014; Muris, 2006). When advising or predicting school related aspects, shedding light on several of these factors and combining them with learning potential might be helpful.

#### 5.4 Overall conclusion

According to this research, we can conclude that our dynamic test procedure was indicating learning potential. Besides, teachers could partly predict learning potential, though appeared to have a hard time dissolving learning potential from performance. An independent device which identifies learning potential, is essential for determining which child should obtain extra attention. This could be done by identifying underperforming children: children with a

high learning potential and low school performance. These children should either benefit from school related intervention or identification of the individual areas of concern.

Learning potential is somewhat related to school performance, yet it does not seem to be an adequate measure to sufficiently support the advice or prediction of school related aspects. Nevertheless, this purpose could be served by combining learning potential with other school related factors.

### 5.5 Limitations and recommendations

In the current study a number of limitations were identified, therefore further refinements to this approach are fundamental. To obtain a developmental perspective, analysing a variety of age groups is needed. However, given the complexity of the procedure, we conducted this study with only one age group in the Netherlands. Therefore, the children tested in this study have an approximately equal cultural background. Since it is possible that these cultural factors have influenced the test results, it can not be generalized to children all over the world. For this reason, future studies should seek to discover how children of different ages and cultures respond to situations such as those outlined in this study.

Our procedure was able to distinguish between children who did receive prompts following the graduated prompt technique and children who did not receive any training. We concluded that the graduated prompt technique enabled to differentiate between these groups. Nevertheless, we can not state that the graduated prompt technique is essential for the change in test scores. It is possible that any other approach has the same, or even more effect. Further research could aid into developing a more refined dynamic approach of giving prompts.

We found that the learning potential estimated by the teacher correlates more with intellectual functioning in a school setting than learning potential as reported by dynamic testing did. It is possible that this discrepancy can be attributed to the fact that intellectual functioning in a school setting is estimated by the teacher, as is learning potential. The total impression of a child and personal preference of the teacher could have functioned as an overarching factor for the estimation of both learning potential and the indicators for intellectual functioning in a school setting. Taken this into account, it is possible that this research overestimated this relation, and that the actual relation between the learning potential estimated by the teacher and intellectual functioning in school setting is smaller than this study reveals.

When analysing the relation between learning potential as reported by dynamic testing and intellectual functioning in a school setting, learning potential is defined by the post-test accuracy score. The generalisation from post-test accuracy scores to the concept of learning potential might be disproportionate. Further research might overcome this by taking more test scores (e.g. pre-test accuracy scores and number of prompts needed) into account when analysing the relation between learning potential and other subjects.

In this research, due to the missing data, the examination of how learning potential is related to intellectual functioning in a school setting, is split into three independent regression analyses. Therefore, it is not clear what the indicators would explain taken into account the remaining indicators. In subsequent studies, this dependent analysis should be performed, provided that there are no missing data.

According to our test results, we can conclude that our dynamic test procedure indicated learning potential, which is important since teachers had difficulty estimating learning potential. The concept of learning potential can be helpful for identifying underperforming children and exploring areas of complication. Learning potential in combination with other school related factors could support the advice about, or prediction of school related aspects.



## 6. Literature

- Akbari, R., & Hosseini, K. (2008). Multiple intelligences and language learning strategies: Investigating possible relations. *System*, 36(2), 141-155.
- Berk, L. E. (2007). *Development through the lifespan*, 4/E. Boston, Massachusetts: Pearson.
- Bosma, T., & Resing, W. (2010). Teacher's appraisal of dynamic assessment outcomes: Recommendations for weak mathematics-performers. *Journal of Cognitive Education and Psychology*, 9(2), 91-115.
- Bosma, T., & Resing, W. C. (2012). Need for instruction: Dynamic testing in education. *European Journal of Special Needs Education*, 27(1), 1-19.
- Bosma, T., & Resing, W.C.M. (2006). Dynamic assessment and a reversal task: A contribution to needs-based assessment. *Educational and Child Psychology*, 23, 81-98.
- Bosma, T., Resing, M., & Wilma, C. (2008). Bridging the gap between diagnostic assessment and classroom practice. *Journal of Cognitive Education and Psychology*, 7(2), 174-198.
- Bradley, R. H., & Corwyn, R. F. (2002). Socioeconomic status and child development. *Annual Review of Psychology*, 53(1), 371-399.
- Caffrey, E., Fuchs, D., & Fuchs, L. S. (2008). The predictive validity of dynamic assessment: A review. *The Journal of Special Education*, 41(4), 254-270.
- Carr, A., Luckin, R., Yuill, N., & Avramides, K. (2013). How mastery and performance goals influence learners' metacognitive help-seeking behaviours when using Ecolab II. In *International handbook of metacognition and learning technologies* (pp. 659-668). New York: Springer.
- Conrad, R. (1976). Ravens standard progressive matrices. *Bulletin of the British psychological society*, 29(6), 184-184.
- Corpus, J. H., & Wormington, S. V. (2014). Profiles of intrinsic and extrinsic motivations in elementary school: A longitudinal analysis. *The Journal of Experimental Education*, 82(4), 480-501.
- Crescentini, C., Seyed-Allaei, S., De Pisapia, N., Jovicich, J., Amati, D., & Shallice, T. (2011). Mechanisms of rule acquisition and rule following in inductive reasoning. *Journal of Neuroscience*, 31(21), 7763-7774.
- Eshel, Y., & Kohavi, R. (2003). Perceived classroom control, self-regulated learning strategies, and academic achievement. *Educational Psychology*, 23(3), 249-260.
- Fabio, R. A. (2005). Dynamic assessment of intelligence is a better reply to adaptive behavior and cognitive plasticity. *The Journal of general psychology*, 132(1), 41-66.

- Ferrara, R. A., Brown, A. L., & Campione, J. C. (1986). Children's learning and transfer of inductive reasoning rules: Studies of proximal development. *Child development*, 57, 1087-1099.
- Gonida, E. N., Karabenick, S. A., Makara, K. A., & Hatzikyriakou, G. A. (2014). Perceived parent goals and student goal orientations as predictors of seeking or not seeking help: Does age matter? *Learning and instruction*, 33, 120-130
- Graczyk, P. A., Domitrovich, C. E., Small, M., & Zins, J. E. (2006). Serving all children: An implementation model framework. *School Psychology Review*, 35(2), 266-274.
- Greiff, S., & Neubert, J. C. (2014). On the relation of complex problem solving, personality, fluid intelligence, and academic achievement. *Learning and Individual Differences*, 36, 37-48.
- Grigorenko, E. L., & Sternberg, R. J. (1998). Dynamic testing. *Psychological Bulletin*, 124(1), 75-111.
- Haavisto, M. L., & Lehto, J. E. (2005). Fluid/spatial and crystallized intelligence in relation to domain-specific working memory: A latent-variable approach. *Learning and Individual Differences*, 15(1), 1-21.
- Hartog, J., Oosterbeek, H., & Teulings, C. (1993). Age, wages and education in the Netherlands. Johnson, P., & Zimmermann, K. F., *Labour Market Implications of European Ageing*. Cambridge, United Kingdom: Cambridge University Press.
- Hayes, B. K., Heit, E., & Swendsen, H. (2010). Inductive reasoning. *Wiley interdisciplinary reviews: Cognitive science*, 1(2), 278-292.
- Haywood, H. C., & Lidz, C. S. (2007). *Dynamic assessment in practice: Clinical and educational applications*. Cambridge, United Kingdom: Cambridge University Press.
- Hessels, M. G. (2009). Estimation of the predictive validity of the HART by means of a dynamic test of geography. *Journal of Cognitive Education and Psychology*, 8(1), 5-21.
- Hojtink, H., Béland, S., & Vermeulen, J. A. (2014). Cognitive diagnostic assessment via Bayesian evaluation of informative diagnostic hypotheses. *Psychological methods*, 19(1), 21-38.
- Hollenberg, J., Lubbe, M. Van der, & Sanders, P. (2017). *Toetsen op school primair onderwijs*. Arnhem, Nederland: Cito B. V.
- Janssen, J., Hop, M., & Wouda, J., (2015). *Wetenschappelijke verantwoording: Rekenen-Wiskunde 3.0 voor groep 4*. Arnhem, Nederland: Cito B.V.

- Jolink, A., Tomesen, M., Hilte, M., Weekers, A., & Engelen, R. (2015). *Wetenschappelijke verantwoording: Begrijpend lezen 3.0 voor groep 4*. Arnhem, Nederland: Cito B.V.
- Kaldenbach, Y. (2006). De WISC-III anno 2006. *Kind & Adolescent Praktijk*, 5(3), 128-136.
- Karabenick, S. A., & Knapp, J. R. (1991). Relationship of academic help seeking to the use of learning strategies and other instrumental achievement behavior in college students. *Journal of Educational Psychology*, 83(2), 221.
- Klauer, K. J., & Phye, G. D. (2008). Inductive reasoning: A training approach. *Review of Educational Research*, 78(1), 85-123.
- Kolakowsky, S. A. (1998). Assessing learning potential in patients with brain injury: Dynamic assessment. *NeuroRehabilitation*, 11(3), 227-238.
- Kossowska, M., & Nęcka, E. (1994). Do it your own way: Cognitive strategies, intelligence, and personality. *Personality and Individual Differences*, 16(1), 33-46.
- Lidz, C. S., & Elliott, J. G. (2000). Introduction in dynamic assessment. In Lidz, C. S., & Elliott, J. G. (Eds.). (2000). *Dynamic assessment: Prevailing models and applications*. 3-17. Sunderland, United Kingdom: University of Sunderland.
- Light, K. E., Reilly, M. A., Behrman, A. L., & Spirduso, W. W. (1996). Reaction times and movement times: Benefits of practice to younger and older adults. *Journal of Aging and Physical Activity*, 4(1), 27-41.
- Liu, O. L. (2009). Evaluation of a learning strategies scale for middle school students. *Journal of Psychoeducational Assessment*, 27(4), 312-322.
- Lu, Z., & Liu, M. (2015). An investigation of Chinese university EFL learner's foreign language reading anxiety, reading strategy use and reading comprehension performance. *Studies in Second Language Learning and Teaching*, 5(1), 65-85.
- McGrew, K. S., & Wendling, B. J. (2010). Cattell–Horn–Carroll cognitive-achievement relations: What we have learned from the past 20 years of research. *Psychology in the Schools*, 47(7), 651-675.
- Misra, K., Grimes, P. W., & Rogers, K. E. (2013). The effects of community social capital on school performance: A spatial approach. *The Journal of Socio-Economics*, 42, 106-111.
- Muris, P. (2006). Relation of attention control and school performance in normal children. *Perceptual and motor skills*, 102(1), 78-80.
- Nelson-Le Gall, S., & Glor-Scheib, S. (1985). Help seeking in elementary classrooms: An observational study. *Contemporary Educational Psychology*, 10(1), 58-71.

- Nijenhuis, J., Willigers, D., Dragt, J., & van der Flier, H. (2016). The effects of language bias and cultural bias estimated using the method of correlated vectors on a large database of IQ comparisons between native Dutch and ethnic minority immigrants from non-Western countries. *Intelligence*, *54*, 117-135.
- Olkun, S. (2003). Comparing computer versus concrete manipulatives in learning 2D geometry. *Journal of Computers in Mathematics and Science Teaching*, *22*(1), 43-56.
- Ordóñez, S., & Sánchez, C. (2016). Knowledge capitalism, globalization, and hegemony: Toward a socio-spatial approach. *World Review of Political Economy*, *7*(1), 4-28.
- Özsoy, G., & Ataman, A. (2009). The effect of metacognitive strategy training on mathematical problem solving achievement. *International Electronic Journal of Elementary Education*, *1*(2), 67-82.
- Raven, J. (2000). The Raven's progressive matrices: change and stability over culture and time. *Cognitive Psychology*, *41*(1), 1-48.
- Resing, W. C., Bakker, M., Pronk, C. M., & Elliott, J. G. (2016). Dynamic testing and transfer: An examination of children's problem-solving strategies. *Learning and Individual Differences*, *49*, 110-119.
- Resing, W. C., Bakker, M., Pronk, C. M., & Elliott, J. G. (2017). Progression paths in children's problem solving: The influence of dynamic testing, initial variability, and working memory. *Journal of Experimental Child Psychology*, *153*, 83-109.
- Resing, W. C., Touw, K. W., Veerbeek, J., & Elliott, J. G. (2017). Progress in the inductive strategy-use of children from different ethnic backgrounds: a study employing dynamic testing. *Educational Psychology*, *37*(2), 173-191.
- Resing, W. C., Tunteler, E., de Jong, F. M., & Bosma, T. (2009). Dynamic testing in indigenous and ethnic minority children. *Learning and Individual Differences*, *19*(4), 445-450.
- Resing, W. C., Xenidou-Dervou, I., Steijn, W. M., & Elliott, J. G. (2012). A "picture" of children's potential for learning: Looking into strategy changes and working memory by dynamic testing. *Learning and Individual Differences*, *22*(1), 144-150.
- Resing, W., & Elliott, J. G. (2011). Dynamic testing with tangible electronics: Measuring children's change in strategy use with a series completion task. *British Journal of Educational Psychology*, *81*(4), 579-605.

- Sfard, A. (1991). On the dual nature of mathematical conceptions: Reflections on processes and objects as different sides of the same coin. *Educational studies in Mathematics*, 22(1), 1-36.
- Shawer, S. F. (2016). Four language skills performance, academic achievement, and learning strategy use in preservice teacher training programs. *TESOL Journal*, 7(2), 262-303.
- Siegel, L. S. (1989). IQ is irrelevant to the definition of learning disabilities. *Journal of Learning Disabilities*, 22(8), 469-478.
- Stegers-Jager, K. M., Cohen-Schotanus, J., & Themmen, A. P. (2012). Motivation, learning strategies, participation and medical school performance. *Medical Education*, 46(7), 678-688.
- Sternberg, R. J. (1985). *Beyond IQ: A triarchic theory of human intelligence*. Cambridge, United Kingdom: Cambridge University Press.
- Sternberg, R. J., & Gardner, M. K. (1983). Unities in inductive reasoning. *Journal of Experimental Psychology: General*, 112(1), 80-115.
- Stevenson, C. E., Bergwerff, C. E., Heiser, W. J., & Resing, W. C. M. (2014). Working memory and dynamic measures of analogical reasoning as predictors of children's math and reading achievement. *Infant and Child Development*, 23, 51–66.
- Tenenbaum, J. B., Griffiths, T. L., & Kemp, C. (2006). Theory-based Bayesian models of inductive learning and reasoning. *Trends in Cognitive Sciences*, 10(7), 309-318.
- Verhaegh, J. (2012). *Assessment and development of cognitive skills using tangible electronic board games: Serious games on the TUI TagTiles*. Eindhoven, Netherlands: Technische Universiteit Eindhoven.
- Verhaegh, J., Fontijn, W., & Hoonhout, J. (2007). *Proceedings of the 1st international conference on Tangible and embedded interaction*. New York, USA: ACM press.
- Vygotsky, L. S. (1980). *Mind in society: The development of higher psychological processes*. Massachusetts, United Kingdom: Harvard university press.