Received: 25 June 2020 | Revised: 1 October 2020 | Accepted: 8 November 2020

DOI: 10.1111/jcal.12512

ARTICLE

Journal of Computer Assisted Learning WILEY

Computerized dynamic testing of children's potential for reasoning by analogy: The role of executive functioning

Bart Vogelaar[1] | Jochanan Veerbeek[2] | Suzanne E. Splinter[3] | Wilma C. M. Resing[1]

[1]Developmental and Educational Psychology, Leiden University, Leiden, The Netherlands

[2]Education and Child Studies, Leiden University, Leiden, The Netherlands

[3]Instructional Psychology and Technology, Catholic University of Louvain, Louvain, Belgium

Correspondence
Dr. Bart Vogelaar, Faculty of Social and Behavioural Sciences, Department of Psychology, Developmental and Educational Psychology Unit, Leiden University, P.O. Box 9555, 2300 RB Leiden, The Netherlands.
Email: b.vogelaar@fsw.leidenuniv.nl

Abstract

This study aimed to investigate children's potential for reasoning by analogy utilizing a newly-developed computerized dynamic test, and the potential differential influence of executive functions (cognitive flexibility, attention, and planning) on static and dynamic measures of analogical reasoning. Participants included 64 children (mean age = 7.55). The study employed a two-session experimental test-training-test design. Based on randomized blocking, half of the children received a graduated prompts training between pre-test and post-test, and the other half did not. Trained children improved more than control children in both their accuracy scores and number of accurately applied transformations from pre-test to post-test. It was further found that cognitive flexibility, attention and planning, is associated with successful solving of analogies. Training children in analogical reasoning seemed to reduce the effect of executive functions. It was also found that that children who were more cognitively flexible needed more prompts during the training.

KEYWORDS

analogical reasoning, computerized dynamic testing, executive functioning, graduated prompts

1 | INTRODUCTION

The ability to reason by analogy is believed to play an important role in everyday learning and problem-solving (Richland & Simms, 2015). This form of inductive reasoning is seen as the cognitive process of comparing information or objects, and, if necessary, contrasting and transforming them in a novel manner (Goswami, 2012). It has been argued that this type of reasoning lies at the core of various cognitive skills and processes, amongst which fluid intelligence (Goswami, 2012; Sternberg, 1985). The extent to which individuals can successfully reason by analogy is, amongst other factors, dependent upon individual characteristics, such as executive functioning (Weatherholt, Harris, Burns, & Clement, 2006).

When measuring children's analogical reasoning abilities, often conventional, also known as static, tests are used. Such tests usually have a single-session format, and consist of children solving tasks individually after short, standardized instructions (Sternberg & Grigorenko, 2002). Although used frequently, they are criticized for focusing predominantly on already developed abilities and prior knowledge, rather than on latent ability, and potential for future learning (Resing, Elliott, & Vogelaar, 2020). As an alternative to static tests, researchers advocate testing dynamically. Dynamic tests are assumed to capture an individual's potential for learning, as they focus on abilities that are not yet fully developed, and include the provision of instruction in the testing procedure (Resing et al., 2020; Sternberg & Grigorenko, 2002). As such, dynamic testing is seen as an approach to measuring children's potential or latent abilities, which is believed to be more equitable than traditional, static tests, for example, due to their ability to compensate for weaknesses in executive functioning (Resing et al., 2020).

Although using dynamic tests for measuring inductive reasoning ability is promising, in practice they are not often used for these purposes, possibly because administration is highly labour-intensive and, thus, expensive (Resing et al., 2020). As a response, some authors advocate using computerized dynamic tests, as a means to gather fine-grained process-oriented data, reduce the labour-intensive nature of testing and allow for provision of individualized instruction (Resing & Elliott, 2011). Therefore, the current study aimed to investigate children's potential for reasoning by analogy, utilizing a newly-developed computerized dynamic test and examine the potential differential influence of executive functions on static and dynamic measures of analogical reasoning.

1.1 | Literature review

1.1.1 | Computerized dynamic testing

There are many different forms of dynamic tests, but in all of them help, in the form of feedback, hints or instruction, is provided to the child when solving tasks (Resing et al., 2020; Sternberg & Grigorenko, 2002). Dynamic tests often have a test-training-test format, and can be combined with graduated prompts training (Campione & Brown, 1987; Poehner & Lantolf, 2013; Resing & Elliott, 2011; Wu, Kuo, & Wang, 2017). As part of this approach, help is provided in the form of prompts, as soon as a child makes a mistake in independent problem-solving. Often based on process models of task-solving (e.g., Resing, 2000; Sternberg, 1985) and highly standardized, prompts are provided hierarchically. Provision of prompts ranges from very general metacognitive, such as activating children's prior knowledge, to cognitive, tailored to the individual test items, to modelling prompts, which consist of step-by-step modelling of the correct solution. Previous studies demonstrated that, irrespective of their levels of improvement or initial ability, children show large individual differences in their need for instruction as measured by graduated prompts training (Resing & Elliott, 2011).

Computerized dynamic tests have been the focus of several studies (De Beer, 2013; Passig, Tzuriel, & Eshel-Kedmi, 2016; Tzuriel & Shamir, 2002), sometimes in combination with a graduated prompts approach (Poehner & Lantolf, 2013; Wu et al., 2017). Studies revealed that children who were administered a computerized dynamic test of analogical reasoning showed equal (Resing, & Elliott, 2011) or even larger (Tzuriel & Shamir, 2002) learning gains and accuracy scores to those who were dynamically tested by a human examiner, whereas being less time-intensive than paper-and-pencil versions (Stevenson, Touw, & Resing, 2011).

Furthermore, computerized dynamic testing is assumed to allow for more fine-grained analysis of the processes that occur during problem-solving (Veerbeek, Vogelaar, Verhaegh, & Resing, 2019). Investigating these processes is believed to provide more insight into variations in children's problem solving behaviour than end-products of reasoning, that is, accuracy scores (e.g., Resing & Elliott, 2011; Tunteler, Pronk, & Resing, 2008; Tzuriel & Galinka, 2000). One process variable often studied concerns the transformations (changes) children can deal with in inductive reasoning tasks (Resing, Touw, Veerbeek, & Elliott, 2017; Tunteler et al., 2008; Vogelaar, Bakker, Hoogeveen, & Resing, 2017).

1.1.2 | Executive functioning

Executive functions are often described as a set of interrelated complex, top-down cognitive processes enabling control of thought and action (Diamond, 2013). In general, these functions are thought to develop with age (Pureza, Gonçalves, Branco, Grassi-Oliveira, & Fonseca, 2013), and children with higher cognitive ability are assumed to have better functions than those with lower ability (Arffa, 2007). In addition, they are believed to play a large role in learning new skills (Diamond, 2013), as well as in classroom learning (Monette, Bigras, & Guay, 2011; Viterbori, Usai, Traverso, & De Franchis, 2015).

In the current study, three different executive functions were investigated: Cognitive flexibility, attention and planning. Cognitive flexibility, also known as switching, is defined as the ability to switch between different tasks, and flexibly adjust thought, perspectives or approaches to a problem to meet changing demands or rules (Diamond, 2013). Flexibility in the process of solving analogies has been linked to success in solving analogies (Thibaut & French, 2016). Research suggests that finding new relations in the analogical reasoning process is dependent upon cognitive flexibility (Glady, Thibaut, French, & Blaye, 2012).

Attention can be defined as the ability to allocate attentional resources amongst different stimuli that occur simultaneously, and is considered to involve a set of different cognitive and executive processes, including, for example, inhibition (Posner, Sheese, Odludaş, & Tang, 2006). Research has shown that in analogy items of the type A:B::C:D focusing attention on the A:B pair led to an increase in performance in analogical reasoning (Glady, French, & Thibaut, 2017). Planning is often defined as the ability to organize behaviour in a sequence of intermediate steps, or operations, to achieve a specific goal (Owen, 1997). Research findings have suggested that planning is correlated with analogical and matrix reasoning, with children with better planning abilities performing better in reasoning (Zook, Davalos, DeLosh, & Davis, 2004).

Research into the relationship between dynamic testing of analogical reasoning and executive functioning remains sparse. In a study involving nine and 10-year-old children, Ropovik (2014) found that a three-factor structure that included inhibition, attention and working memory predicted children's performance on figural visual–spatial analogy items after a dynamic training best. Other studies have found that visuo-spatial working memory (Resing, 2013) and metacognition (Vogelaar et al., 2017) were related to improvements in analogical reasoning after training. In a study by Vogelaar, Resing, Stad, and Sweijen (2019) it was, furthermore, found that planning predicted the number of prompts children needed in the first training session, but not the second.

1.2 | The current study

The current study aimed to broaden the existing knowledge regarding the relationship between analogical reasoning and executive functioning by examining children's potential for solving analogies, using a newly developed computerized dynamic test of analogical reasoning. As dynamic testing is suggested to compensate for weaknesses in these functions (Vogelaar et al., 2017, Vogelaar et al., 2019), it was examined whether they would be differentially related to static and dynamic measures of analogical reasoning. In the test utilized in the current study, children had to construct their own answers, which was believed to provide insight into the process of task-solving.

Our first aim was to analyse the suitability of our newly developed computerized dynamic test for seven and eight your old children with various intelligence ranges. The first research question concerned the effectiveness of the dynamic test. We focused on (a) accuracy scores, and (b) accurately applied transformations. For both measures, it was expected that trained children would show more improvement from pre-test to post-test than untrained children (Resing et al., 2017; Resing & Elliott, 2011).

Secondly, the relationship between executive functions and static and dynamic measures of analogical reasoning was investigated. Pre-test as well as post-test accuracy and accurately applied transformations of the untrained children were considered static measures, whereas these post-test measures of the dynamically trained children were considered to be dynamic. It was expected that cognitive flexibility, attention and planning would significantly predict the static measures pre-test accuracy and accurately applied transformations (Glady et al., 2017; Hummel & Holyoak, 2003; Weatherholt et al., 2006; Zook et al., 2004). Stronger executive functions were hypothesized to predict higher scores at pre-test.

Children's post-test measures (accuracy and accurately applied transformations) were analysed separately for trained and untrained children. For the post-test measures of the children in the control condition, who did not receive training, similar predictive relationships were expected as for the pre-test. In relation to the children who were dynamically trained, however, due to the compensating nature of dynamic testing, it was expected that the three executive functions would no longer predict the two post-test measures (Vogelaar et al., 2017, Vogelaar et al., 2019).

Finally, the relationship between executive functions and instructional needs was examined. It was expected that cognitive flexibility, attention and planning would significantly predict children's need for instructions (Vogelaar et al., 2019). For this research question, instructional needs were analysed by focusing on the total number of prompts, as well as on the number of metacognitive, cognitive and modelling prompts separately.

2 | METHOD

2.1 | Participants

In total, 64 seven and 8-year-old children (30 boys, 34 girls) participated in the study, with a mean age of 7.55 (SD = 0.50). The children were recruited from seven primary schools in the western part of the Netherlands, and attended schools in middle class neighbourhoods. The intelligence range of the participating children was found to be large, with Raven Standard Progressive Matrices (RSPM, Raven, Raven, & Court, 2000) percentile scores between 10 and 95 (M = 72.11, SD = 25.80). Prior to participation in the study, approval from the university's board of ethics, and written informed consent for children's participation from all parents as well as the schools' headmasters were obtained.

2.2 | Design and procedure

The study used a two-session experimental test-training-test design with two conditions (see Table 1). In the first session, all participants were administered the RSPM (Raven et al., 2000). A randomized blocking procedure based on initial reasoning ability, as measured with the RSPM, allocated the participants to either a training or a control condition. Pairs of children with similar scores were randomly assigned to the training or the control condition.

The second session consisted of administration of the Trail Making Test (TMT) and Tower of Hanoi (ToH). In the third session, all participants were administered the pre-test of the dynamic test. In the third and fourth session, participants in the training condition received a graduated prompts training. Participants in the control condition did not receive training or practice opportunities, but solved control tasks consisting of dot-to-dot tasks and mazes. In the fifth and final session, all participants completed the post-test.

2.3 | Materials

2.3.1 | Raven's standard progressive matrices

The RSPM (Raven et al., 2000) includes 60 multiple choice items requiring participants to identify the missing part of a pattern. The RSPM measures fluid intelligence, and has been found to have a high level of internal consistency (split-half r = 0.91; Raven et al., 2000). In the present study, the test was administered on paper and used as an indication of a child's initial inductive reasoning ability.

2.3.2 | Trail making test

The TMT was used as a measure of children's attention abilities and cognitive flexibility. Children completed the Psychology Experiment Building Language (PEBL) version (Mueller & Piper, 2014), which is a computerized version of the Halstead-Reitan Neuropsychological Test Battery (Reitan, 1955). This version of the TMT involves clicking on circles with either numbers that need to be connected in numerical order (part A: 1–2–3-4-5) or numbers and letters that need to be connected in alternating order (part B: 1-A-2-B-3-C). Part A can be seen as an index of visual attention,

TABLE 1 Overview of the study's design

	Raven	Trail making test	Tower of Hanoi	Pre-test	Training 1	Training 2	Post-test
Training (*n* = 32)	x	x	x	x	x	x	x
Control (*n* = 32)	x	x	x	x	-	-	x

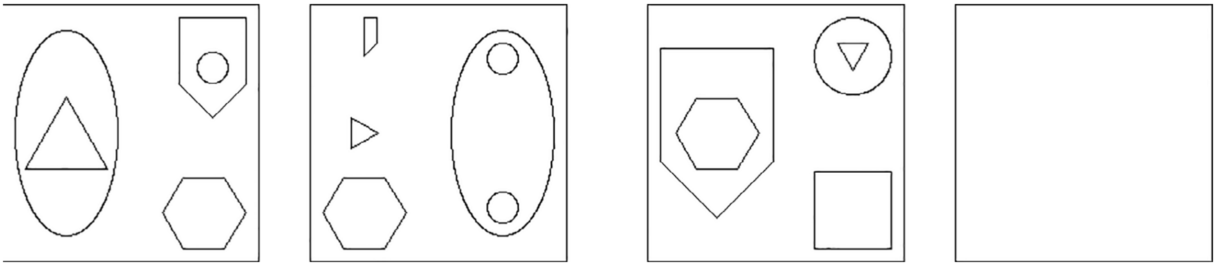

FIGURE 1 Example of a difficult analogy item

whereas part B is thought to reflect cognitive flexibility (Piper et al., 2012). Test–retest reliabilities of r = 0.74 for completion time part A and r = 0.61 for completion time part B have been reported (Piper et al., 2015).

2.3.3 | Tower of Hanoi

The digital PEBL version of the ToH was administered as a measure of children's planning abilities (Mueller & Piper, 2014). The ToH, originally developed by Kotovsky, Hayes, and Simon (1985), involves moving disks of differing sizes to match a desired goal in a minimum number of moves and as quickly as possible. Children received the instructions that larger disks could not be placed on smaller disks and that only one disk could be moved at a time. Low to satisfactory test–retest reliabilities have been found (Ahonniska, Ahonen, Aro, Tolvanen, & Lyytinen, 2000).

2.3.4 | Computerized dynamic test of analogical reasoning

The computerized dynamic test consisted of open-ended visuospatial geometric analogies of the type A:B::C:? (see Figure 1 for an example item). The items were developed by Vogelaar, Sweijen, and Resing (2019) for use with average and high-ability children, and were based on item sets originally constructed by Tunteler & Colleagues (2008). Each analogy item contained a maximum of six possible basic geometrical shapes (circles, squares, triangles, ellipses, hexagons and pentagons) and two to fourteen possible transformations (one of the following changes to an element: adding or subtracting an element, changing size, changing position, rotating or halving).

Pre-test and post-test

The pre and post-test were constructed as parallel versions, and included fifteen items that increased in difficulty. Participants were provided with a short, general instruction prior to the test session, but did not receive any additional help.

Graduated prompts training

The training consisted of two short sessions containing six new analogy items each. The items were presented on a laptop, and help, verbalized by the examiners, was provided on the basis of a graduated prompts approach (Resing, 2000; Campione & Brown, 1987). The order of the prompts provided was based on previous process models of solving analogical reasoning (Resing, 2000; Sternberg, 1985). Prompts were given if a child provided an incorrect answer, and were provided in a hierarchical fashion, starting with two general, metacognitive prompts and ending with three item-specific cognitive prompts. If the participant was not able to provide the correct answer after the last cognitive prompt, the final prompt consisted of modelling of the correct answer. The final step consisted of the examiner asking the child for an explanation of why they thought their given answer was correct.

Computerization of the dynamic test

The analogy items were adapted for use on a laptop, so that each child could construct their solutions to the analogies, utilising a computer mouse to drag and drop the shapes they thought they needed into the fourth empty box. For each item to be solved, children were provided with all the different shapes that were used in the construction of the analogies. They could then use a range of buttons to apply the transformations they thought they needed to each individual shape (see Figure 2), reset their solution, in case of a potential mistake and continue to the next item.

The computerized test version was made in Eclipse IDE (http://www.eclipse.org/ide) using the Java language. The possible transformations for each figure were provided with a unique three digit identifier, in which the first digit represented the shape, the second digit described whether the shape was whole, half, or half mirrored and the third digit described the rotation. To allow for three different sizes, and the ability to score three differently sized figures on the same position, the answering window was built up using three layers, one for each size. Additionally, the window was divided into nine zones, to

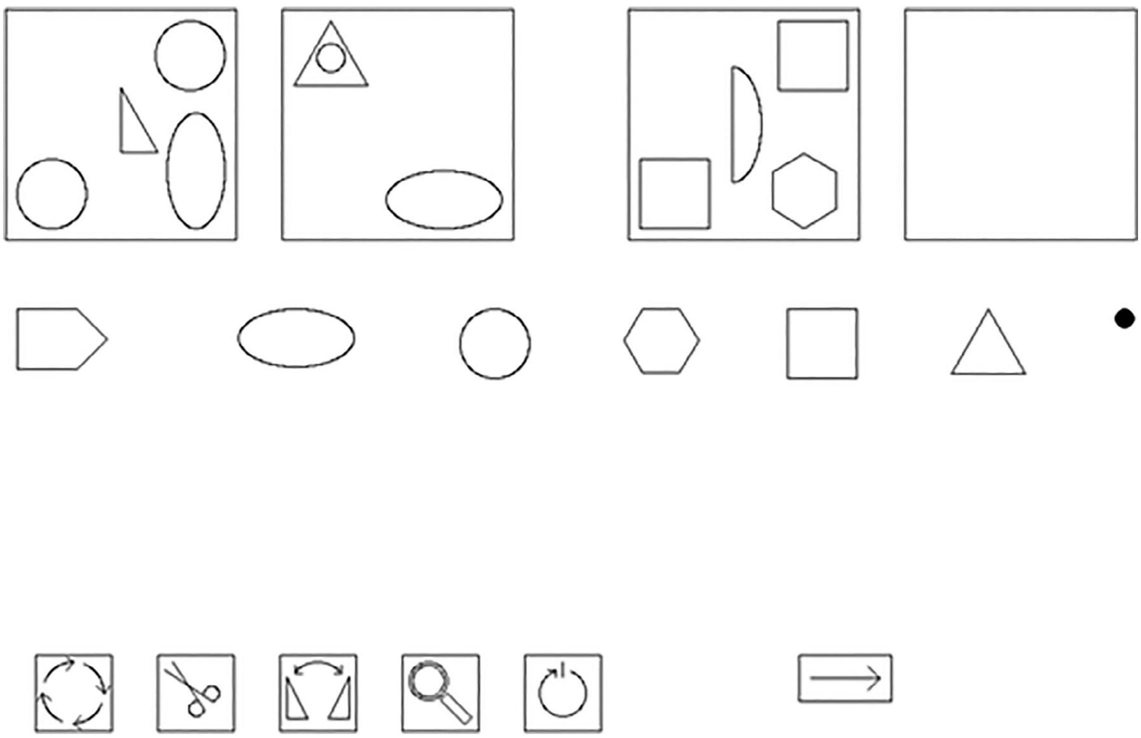

FIGURE 2 Screenshot of an item of the computerized test, including answer possibilities for each individual analogy item, as well as the buttons children could use for applying transformations, resetting their solution, and continuing to the next item

facilitate the scoring of the placement of the figures in the correct position. The programme included automated scoring and for the placement of each figure it recorded the timestamp, place, size and its identifier for the shape and transformations into log files.

2.4 | Scoring

2.4.1 | TMT

Two measures of the TMT were used in our analyses: completion times and accuracy. The total time and accuracy scores for parts A and B and for each trial were computed by the PEBL. Completion times, provided in seconds, for parts A and B, the total time children needed to complete the trials in each part, were thought to reflect efficiency scores and have traditionally been used as indices of attention (part A) and cognitive flexibility (part B) (Piper et al., 2012). Mean scores of the accuracy for the items pertaining to parts A and B were also computed. These scores represented the minimum number of clicks necessary to complete each item divided by the number made. Lower completion times and higher accuracy scores correspond with higher attention abilities and cognitive flexibility.

2.4.2 | ToH

For the ToH, completion times and efficiency were used. These two measures were computed by the PEBL. Completion time, provided in seconds, refers to the time children needed to complete all trials. An efficiency score, defined by the minimum number of moves necessary to solve all trials divided by the number of moves taken to complete them, was also calculated. Lower completion times and higher efficiency scores indicate higher planning abilities.

2.4.3 | Pre-test and post-test measures of analogical reasoning

For the pre-test and the post-test both the number of correct items as well as the number of accurately applied transformations was utilized. These scores were automatically calculated by the programme, and provided in automated log files.

3 | RESULTS

3.1 | Psychometric properties

First, the psychometric properties of the dynamic test were analysed. *p*-values were calculated to assess the difficulty level of each item. These values were defined as the proportion of participants that solved the items accurately, and as the proportions of accurately applied transformations of each item. For the pre-test, the *p*-values for number of accurately solved items ranged from .09 to .58 and for the post-test from .34 to .81 (dynamic testing condition) and .13 to .62 (control condition). The *p*-values for the number of accurately

TABLE 2 Basis statistics for age, Raven percentiles, pre-test and post-test accuracy and transformations, instructional needs, and executive function measures divided by condition

	Training (n = 32)		Control (n = 32)		Total (n = 64)	
	M	*SD*	*M*	*SD*	*M*	*SD*
Age	7.56	0.50	7.53	0.51	7.55	0.50
Raven percentiles	69.69	25.08	74.53	26.68	72.11	25.80
Pre-test accuracy	4.59	4.51	3.94	3.87	4.27	4.18
Pre-test transformations	46.94	26.10	44.28	26.60	45.61	26.18
Post-test accuracy	10.03	4.96	7.19	6.13	8.61	5.71
Post-test transformations	63.97	25.19	50.13	32.43	57.05	29.64
Metacognitive prompts	11.69	8.37				
Cognitive prompts	7.75	8.56				
Modelling prompts	2.97	3.93				
Total prompts	21.50	19.63				
Planning completion time	377.36	207.89	386.88	233.92	382.12	219.58
Planning efficiency	81.38	27.16	83.66	29.83	82.52	28.32
Attention completion time	133.17	35.67	125.16	23.52	129.16	30.24
Attention accuracy	0.94	0.04	0.92	0.05	0.93	0.04
Cogn. Flexibility completion time	200.72	63.53	186.42	49.00	193.56	56.74
Cogn. Flexibility accuracy	0.88	0.05	0.88	0.08	0.88	0.07

TABLE 3 Results of the repeated measures MANOVA on accuracy scores and number of accurately applied transformations

	Wilks' λ	*F*	*p*	η_p^2
Multivariate effects				
Session	0.42	41.50	<.001	0.58
Session × condition	0.88	4.13	.021	0.12
Univariate effects				
Accuracy				
Session		83.95	<.001	0.58
Session × condition		5.32	.024	0.08
Accurately applied transformations				
Session		34.17	<.001	0.36
Session × condition		8.17	.006	0.12

applied transformations for the pre-test ranged from .44 to .71, and for the post-test from .68 to .86 (dynamic testing condition) and .53 to .67 (control condition). Item-total correlations appeared to be moderate to high, with values for the pre-test ranging from r_{it} = 0.37 to r_{it} = 0.77, and for the post-test from r_{it} = 0.42 to r_{it} = 0.89 (dynamic testing condition) and from r_{it} = 0.47 to r_{it} = 0.96 (control condition). The full range of *p*-values and item-total correlations can be found in Table A1 in the Appendix.

Furthermore, the internal consistencies were high: α = 0.90 for the pre-test and α = 0.94 (dynamic testing condition) and α = 0.97 (control condition) for the post-test. As expected, the test–retest reliability was significantly higher for the control condition (r = 0.86, p < .001) than for the dynamic testing condition (r = 0.63, p < .001), as indicated by Fisher's *r* to *z* transformation (z = 2.06, p = .039).

3.2 | Initial group comparisons

Before conducting our analyses, possible differences between the two conditions in age, initial reasoning performance (Raven percentile score), pre-test measures (accuracy and accurately applied transformations) and executive functions were examined using a one-way MANOVA. The multivariate results revealed no significant differences between trained and control children, Wilks' λ = 0.86, F(10, 53) = 0.83, p = .599, η_p^2 = 0.14. Means and standard deviations of the variables examined are presented in Table 2.

3.3 | Training effects

Next, it was analysed whether the trained children improved more in analogy problem-solving than the control children. Differences between pre and post-test were investigated using a repeated measures MANOVA with accuracy scores and the number of accurately applied transformations as dependent variables, Condition (training/control) as the between-subjects factor, and Session (pre-test/post-test) as the within-subjects factor.

The outcomes of the analysis are presented in Table 3. A significant main effect of Session indicated that, overall, all groups of children improved from pre to post-test on the number of correctly solved analogies and the number of accurately applied transformations. Further, a significant Session × Condition effect indicated a significant difference between the two conditions in their level of improvement in accuracy scores and accurately applied transformations. Follow-up univariate analyses and inspection of mean scores (see Table 2 and Figure 3) revealed, as expected, that trained children

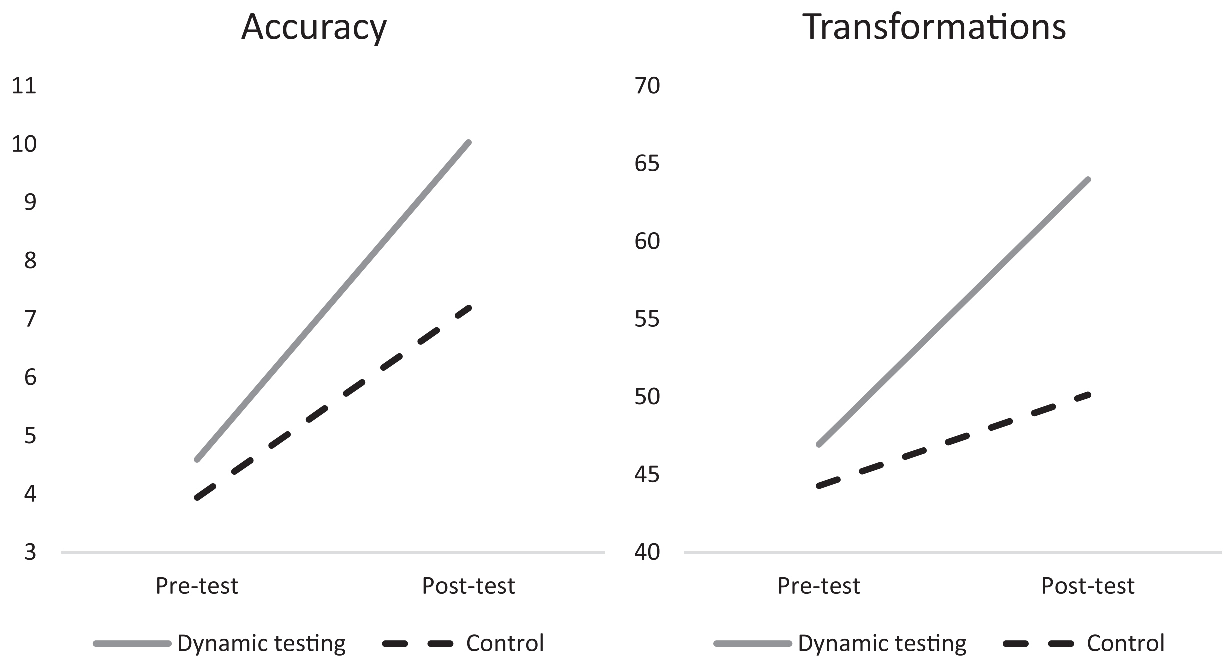

FIGURE 3 Accuracy scores (left) and accurately applied transformations (right) for each condition

TABLE 4 Pearson product–moment correlations between pre-test and post-test (divided by condition), instructional needs, and executive function measures

			Post-test accuracy		Post-test transformations		Instructional needs			
	Pre-test accuracy	Pre-test transformations	Training	Control	Training	Control	Meta cogn.	Cogn.	Mod.	Total
Planning compl. Time	−0.28*	−0.34**	−0.36*	−0.22	−0.31	−0.32	0.48**	0.45*	0.40*	0.45*
Planning efficiency	−0.27*	−0.32**	−0.28	−0.26	−0.22	−0.33	0.44*	0.37*	0.29	0.36*
Attention compl. time	−0.19	−0.32*	−0.33	−0.33	−0.36*	−0.41*	0.27	0.30	0.36*	0.33
Attention accuracy	0.27*	0.33**	0.19	0.45**	0.16	0.53**	−0.10	−0.09	−0.14	−0.10
Cogn. flexibility compl. time	−0.30*	−0.38**	−0.38*	−0.43*	−0.41*	−0.43*	0.43*	0.45**	0.53**	0.49**
Cogn. flexibility accuracy	0.49**	0.60**	0.49**	0.67**	0.39*	0.73**	−0.47**	−0.41*	−0.37*	−0.43*

**p* < .05. ***p* < .01.

improved more than control children in both their accuracy scores and number of accurately applied transformations.

3.4 | Relationship between executive functions and measures of analogical reasoning

Prior to performing linear regression analyses to examine the relationship between executive functioning and measures of analogical reasoning, Pearson correlations were calculated to explore the relationships between these variables (see Table 4). Overall, it was found that the executive functioning measures correlated significantly with the analogical reasoning measures. Each executive function examined demonstrated different strengths of relationships, but all the relationships found were in the hypothesized direction: stronger executive functions corresponded with higher accuracy scores and a larger number of accurately applied transformations at the pre-test and post-test, and fewer prompts. As expected, relationships between the static post-test measures (of untrained children) and the executive functions were, in general, stronger than for the dynamic post-test measures (of trained children).

The fact that the completion times and efficiency measures demonstrated differential correlation patterns supported the view that they measure different aspects of executive functioning (Van der

TABLE 5 Multiple regression analysis predicting pre-test and post-test (divided by condition) by executive function measures

	Pre-test				Post-test training				Post-test control			
	Accuracy		Transformations		Accuracy		Transformations		Accuracy		Transformations	
Variable	*B (SE)*	β	*B (SE)*	β	*B (SE)*	β	*B (SE)*	β	*B (SE)*	β	*B (SE)*	β
Constant	−17.83 *(11.23)*		−105.54 *(63.78)*		−29.17 *(21.66)*		−94.45 *(115.88)*		−30.07 *(23.11)*		−169.95 *(112.72)*	
Planning compl. time	−0.00 *(0.01)*	−0.12	0.01 *(0.03)*	−0.06	−0.01 *(0.01)*	−0.28	−0.02 *(0.05)*	−0.16	0.01 *(0.01)*	0.21	−0.00 *(0.04)*	−0.01
Planning efficiency	0.01 *(0.04)*	−0.05	0.03 *(0.21)*	0.03	0.02 *(0.06)*	0.12	−0.07 *(0.33)*	0.08	−0.00 *(0.07)*	−0.02	0.16 *(0.33)*	0.14
Attention compl. time	0.02 *(0.03)*	0.15	−0.03 *(0.15)*	−0.03	0.02 *(0.05)*	0.14	0.02 *(0.25)*	0.03	−0.02 *(0.05)*	−0.07	−0.20 *(0.25)*	−0.14
Attention accuracy	−0.73 *(13.82)*	−0.01	−20.16 *(78.47)*	−0.03	9.84 *(24.85)*	0.08	61.57 *(132.92)*	0.09	−1.38 *(27.86)*	−0.01	−13.35 *(135.87)*	−0.02
Cogn. flex. compl. time	−0.02 *(0.03)*	−0.24	−0.08 *(0.08)*	−0.17	−0.03 *(0.03)*	−0.34	−0.14 *(0.13)*	−0.34	−0.02 *(0.02)*	−0.17	−0.04 *(0.12)*	−0.06
Cogn. flex. accuracy	26.93 *(10.02)*	0.44**	213.31 *(56.90)*	0.56**	38.00 *(19.27)*	0.40	143.27 *(103.08)*	0.30	48.55 *(17.73)*	0.65*	285.14 *(86.50)*	0.72**
R^2	0.279		0.407		0.360		0.290		0.483		0.560	
F	3.67**		6.52***		2.35		1.70		3.89**		5.31**	

$^{*}p < .05$. $^{**}p < .01$. $^{***}p < .001$.

Sluis, De Jong, & Van der Leij, 2007), and provide support for utilizing both types of measures in research.

3.4.1 | Pre-test

Two linear regression analyses were performed to investigate the potential influence of executive functions on pre-test accuracy and accurately applied transformations. All executive function measures were simultaneously entered into the regression models. The regression model included the main effects of Planning (completion time and efficiency), Attention (completion time and accuracy) and Cognitive flexibility (completion time and accuracy). The results are displayed in Table 5. In the first model, only accuracy in cognitive flexibility was found to significantly predict pre-test accuracy (b = 26.93, p = .009). The first model explained 27.9% of the variance in pre-test accuracy. In the second model, also only accuracy in cognitive flexibility was found to significantly predict pre-test transformations (b = 213.31, p < .001). The second model explained 40.7% of the variance in pre-test transformations. Children with higher cognitive flexibility scores obtained higher accuracy scores and applied a larger number of transformations accurately at pre-test.

3.4.2 | Post-test

For each of the two post-test measures, accuracy and accurately applied transformations, we conducted two separate linear regression analyses, one for each condition, to examine the effects of executive functions on static and dynamic measures of analogical reasoning. For all analyses, the models used included the main effects of the six executive function measures (efficiency scores and completion times). The results of these analyses are also displayed in Table 5. For the dynamic testing condition, the models for accuracy and accurately applied transformations were both non-significant. The first model explained 36.0% of the variance in post-test accuracy and the second model explained 29.0% of the variance in accurately applied transformations.

Regarding the control condition, similar to the pre-test measures, only the cognitive flexibility accuracy score was found to significantly predict post-test accuracy (b = 48.55, p = .011), and transformations (b = 285.14, p = .003) of the untrained children. The first model explained 48.3% of the variance in post-test accuracy and the second model explained 56.0% of the variance in accurately applied transformations. Children with higher cognitive flexibility scores had higher accuracy scores and applied a larger number of transformations accurately at post-test.

3.4.3 | Instructional needs

Finally, we performed four separate linear regression analyses to investigate the effects of executive functions on the number of

TABLE 6 Multiple regression analysis predicting instructional needs by executive function measures

Variable	Instructional needs: Metacognitive B (SE)	β	Cognitive B (SE)	β	Modelling B (SE)	β	Total B (SE)	β
Constant	60.09 *(32.40)*		54.13 *(34.96)*		26.37 *(15.80)*		135.17 *(78.16)*	
Planning compl. time	0.01 *(0.01)*	0.33	0.02 *(0.02)*	0.40	0.01 *(0.01)*	0.34	0.03 *(0.03)*	0.36
Planning efficiency	0.03 *(0.09)*	0.10	−0.01 *(0.10)*	−0.03	−0.01 *(0.05)*	−0.06	−0.02 *(0.22)*	−0.03
Attention compl. time	−0.15 *(0.07)*	−0.65*	−0.14 *(0.08)*	−0.60	−0.06 *(0.04)*	−0.56	−0.31 *(0.17)*	−0.56
Attention accuracy	−14.23 *(37.17)*	−0.07	−16.37 *(40.10)*	−0.07	−15.88 *(18.12)*	−0.15	−38.47 *(89.65)*	−0.08
Cogn. flex. compl. time	0.10 *(0.04)*	0.76*	0.10 *(0.04)*	0.74*	0.05 *(0.02)*	0.85**	0.24 *(0.09)*	0.76*
Cogn. flex. accuracy	−48.03 *(28.83)*	−0.30	−42.49 *(31.01)*	−0.26	−14.20 *(14.06)*	−0.19	−107.78 *(69.52)*	−0.29
R^2	0.497		0.441		0.458		0.468	
F	4.12**		3.29*		3.52*		3.67**	

*$p < .05$. **$p < .01$.

metacognitive, cognitive, modelling and total number of prompts children needed. Each model included the main effects of planning (completion time and efficiency), attention (completion time and accuracy) and cognitive flexibility (completion time and accuracy). Results are depicted in Table 6. For all four analyses Cognitive flexibility completion time was a significant predictor (b = 0.10, p = .012; b = 0.10, p = .019; b = 0.05, p = .007; b = 0.24, p = .014, for metacognitive, cognitive, modelling and total prompts, respectively). For the first model with metacognitive prompts, Attention completion time was also a significant predictor (b = −0.15, p = .042). The first model explained 49.7% of the variance in metacognitive prompts, the second model explained 44.1% of the variance in cognitive prompts, the third model explained 45.8% of the variance in modelling prompts, and the fourth model explained 46.8% of the variance in total prompts. For all four models, these findings suggested that children who spent longer on solving the cognitive flexibility tasks needed more prompts.

4 | DISCUSSION

The current study investigated whether a computerized dynamic test of analogical reasoning could be used to provide insight into the potential for reasoning by analogy of 7- and 8-year-old children. Overall, in accordance with other studies into computerized dynamic testing (Passig et al., 2016; Tzuriel & Shamir, 2002; Wu et al., 2017) our dynamic test seemed effective in unveiling the participants' potential for solving analogies. As compared to control children, trained children demonstrated larger improvements in accuracy scores and accurately applied transformations. An analysis of the mean scores, standard deviations and p-values further suggested that the items used in the computerized dynamic test of analogical reasoning were sufficiently complex for children of 7 and 8 years old.

The second focus of the current study concerned the potential relationship between executive functioning and static and dynamic measure of analogy problem-solving. In general, the findings revealed that pre-test and post-test scores of the untrained children were associated more strongly with all three executive functions than the post-test scores of the trained children. All associations indicated that stronger executive functions were associated with higher scores in analogical reasoning at pre-test and post-test. Linear regression analyses further suggested, in line with the hypotheses, that both the accuracy scores and the accurately applied transformations at pre-test measures of all children, and post-test of the untrained children, but not of those in the experimental condition, could be predicted significantly by children's executive functioning. The only significant predictor, however, was accuracy in cognitive flexibility.

In relation to children's instructional needs, it was found that executive functioning measures, with the exception of attention accuracy, were significantly associated with the number of metacognitive, cognitive, modelling and total number of prompts children needed during training. Furthermore, the number of prompts children needed in total, as well as when investigating the different type of prompts separately (metacognitive, cognitive and modelling), could be significantly predicted by cognitive flexibility.

All in all, as was hypothesized, these findings provided us with preliminary support that computerized dynamic testing of analogical reasoning can compensate for weaknesses in executive functioning, as was also found in previous studies utilizing paper-and-pencil dynamic tests (Vogelaar et al., 2017; Vogelaar et al., 2019). In addition, in accordance with our hypothesis, children with weaker executive functions were found to need more prompts in total, and more specific prompts than those children with stronger executive functions.

In interpreting these results, the multifaceted nature of executive functions needs to be taken into account, which has been argued to lead to differential correlational patterns in other studies (Miyake & Friedman, 2012). In relation to this, both task impurity problems (Friedman et al., 2008), and the developmental nature of executive functions in children (Gathercole, Pickering, Ambridge, & Wearing, 2004) make assessment of executive functioning in these children challenging. Moreover, the fact that completion times and efficiency scores demonstrated differential correlational patterns support the notion that it is worthwhile investigating both efficiency and time scores.

Our study also encountered some limitations. As a consequence of the design, trained children were exposed more to solving analogies than control children, who completed control tasks during training, which could in part have influenced the larger improvements found in the group of trained children. Nevertheless, previous studies that employed a design with two control conditions, one of which practiced with the same items as the trained children, and one of which provided with alternative tasks; however, indicated that the children who were trained demonstrate improvements beyond and above those in the two control conditions (Resing, 1993; Stevenson et al., 2013). Although highly labour-intensive, due to the larger number of participants necessary for such a design, future studies utilizing our computerized dynamic test could employ a similar design with two control conditions to further validate the effectiveness of the computerized dynamic test. In addition, the current study employed a relatively low number of participants, which could have resulted in power loss. Future studies should therefore aim at employing larger numbers of participants, preferably with a larger age range.

Future studies could, in addition, focus on further computerization and adaptability of the test. In the current version, the training items were provided on the laptop screen, but the prompts were verbalized by the examiners. Future studies should aim at having the prompts provided by the laptop, improving standardization. In addition, more items of varying difficulties could be constructed to investigate more closely the type of, and individual differences in, solving behaviour that children of varying ages demonstrate while solving items of varying difficulty level. Such alterations to the test would allow for group-administered and adaptive dynamic testing for children of a wider age range, which would lead to the process of dynamic testing being less labour-intensive, and more cost-effective. As such, such changes would enhance its attractiveness for usage in educational settings. As the current study was a pilot study to investigate the usefulness and suitability of the computerized dynamic test, we only focused on accuracy scores and accurately applied transformations. In future studies, more process variables could be investigated, such as completion time, but also the type of transformations children were able to apply accurately.

5 | CONCLUSION

In conclusion, the current study demonstrated that computerized dynamic tests are an excellent means for measuring the potential for solving analogies of children. In relation to the role of executive functions in children's learning to solve analogies, it seems that cognitive flexibility especially, and to a lesser extent attention and planning, is associated with successful solving of analogies. Training children in analogical reasoning seems to reduce the effect of cognitive flexibility. In terms of the instructions children need when learning to solve analogy items, it seems that those who are more cognitively flexible need fewer instructions when learning to solve analogy items. In addition, children who have better attention skills seemingly need fewer instructions in learning how to solve new analogy items.

The advantages of computerized dynamic testing of analogical reasoning, including automated scoring and the possibility of group-administered and adaptive testing, make it a useful instrument for identification purposes. More importantly, these tests can offer an abundance of information about children's (changes in) problem-solving behaviour, and the type and number of instructions they need when solving new tasks, offering insights relevant for individualized education plans. Such plans are of the utmost importance for providing suitable education to children, who, as has been demonstrated in the current study, demonstrate large individual differences in their ability, subsequent improvement in learning and the instructions they benefit from.

ACKNOWLEDGEMENTS

This research was funded by Stichting Abbas, grant number Abbas Ontwikkelingssubsidie 2017.

CONFLICTS OF INTEREST

The authors declare no conflict of interest.

DATA AVAILABILITY STATEMENT

Data are available upon request

ORCID

Bart Vogelaar https://orcid.org/0000-0002-5131-2480
Jochanan Veerbeek https://orcid.org/0000-0001-9280-9066
Wilma C. M. Resing https://orcid.org/0000-0003-3864-4517

REFERENCES

Ahonniska, J., Ahonen, T., Aro, T., Tolvanen, A., & Lyytinen, H. (2000). Repeated assessment of the tower of Hanoi test: Reliability and age effects. *Assessment*, *7*, 297–310. https://doi.org/10.1177/107319110000700308

Arffa, S. (2007). The relationship of intelligence to executive function and non-executive function measures in a sample of average, above average, and gifted youth. *Archives of Clinical Neuropsychology*, *22*, 969–978. https://doi.org/10.1016/j.acn.2007.08.001

Campione, J. C., & Brown, A. L. (1987). Linking dynamic assessment with school achievement. In C. S. Lidz (Ed.), *Dynamic assessment: An interactional approach to evaluating learning potential* (pp. 82–109). New York, NY: Guilford Press.

De Beer, M. (2013). The learning potential computerised adaptive test in South Africa. In S. Laher & K. Cockcroft (Eds.), *Psychological assessment in South Africa: Research and applications* (pp. 137–157). Johannesburg, South Africa: Wilts University Press.

Diamond, A. (2013). Executive functions. *Annual Review of Psychology*, *64*, 135–168. https://doi.org/10.1146/annurev-psych-113011-143750

Friedman, N. P., Miyake, A., Young, S. E., DeFries, J. C., Corley, R. P., & Hewitt, J. K. (2008). Individual differences in executive functions are almost entirely genetic in origin. *Journal of Experimental Psychology: General*, *137*, 201–225. https://doi.org/10.1037/0096-3445.137.2.201

Gathercole, S. E., Pickering, S. J., Ambridge, B., & Wearing, H. (2004). The structure of working memory from 4 to 15 years of age. *Developmental Psychology*, *40*, 177–190. https://doi.org/10.1037/0012-1649.40.2.177

Glady, Y., French, R. M., & Thibaut, J. P. (2017). Children's failure in analogical reasoning tasks: A problem of focus of attention and information integration? *Frontiers in Psychology*, *8*, 1–13. https://doi.org/10.3389/fpsyg.2017.00707

Glady, Y., Thibaut, J. P., French, R. M., & Blaye, A. (2012). Explaining children's failure in analogy making tasks: A problem of focus of attention? In N. Miyake, D. Peebles, & R. P. Cooper (Eds.), *Proceedings of the thirty-fourth annual meeting of the cognitive science meeting* (pp. 384–389). Austin, TX: Cognitive Science Society.

Goswami, U. C. (2012). Analogical reasoning by young children. In N. M. Seel (Ed.), *Encyclopedia of the sciences of learning* (pp. 225–228). New York, NY: Springer. https://doi.org/10.1007/978-1-4419-1428-6_993

Hummel, J. E., & Holyoak, K. J. (2003). A symbolic-connectionist theory of relational inference and generalization. *Psychological Review*, *110*, 220–264. https://doi.org/10.1037/0033-295x.110.2.220

Kotovsky, K., Hayes, J. R., & Simon, H. A. (1985). Why are some problems hard? Evidence from tower of Hanoi. *Cognitive Psychology*, *17*(2), 248–294. https://doi.org/10.1016/0010-0285(85)90009-X

Miyake, A., & Friedman, N. P. (2012). The nature and organization of individual differences in executive functions: Four general conclusions. *Current Directions in Psychological Science*, *21*, 8–14. https://doi.org/10.1177/0963721411429458

Monette, S., Bigras, M., & Guay, M. C. (2011). The role of the executive functions in school achievement at the end of grade 1. *Journal of Experimental Child Psychology*, *109*, 158–173. https://doi.org/10.1016/j.jecp.2011.01.008

Mueller, S. T., & Piper, B. J. (2014). The psychology experiment building language (PEBL) and PEBL test battery. *Journal of Neuroscience Methods*, *222*, 250–259. https://doi.org/10.1016/j.jneumeth.2013.10.024

Owen, A. M. (1997). Cognitive planning in humans: Neuropsychological, neuroanatomical and neuropharmacological perspectives. *Progress in Neurobiology*, *53*, 431–450. https://doi.org/10.1016/S0301-0082(97)00042-7

Passig, D., Tzuriel, D., & Eshel-Kedmi, G. (2016). Improving children's cognitive modifiability by dynamic assessment in 3D immersive virtual reality environments. *Computers & Education*, *95*, 296–308. https://doi.org/10.1016/j.compedu.2016.01.009

Piper, B. J., Li, V., Eiwaz, M. A., Kobel, Y. V., Benice, T. S., Chu, A. M., … Mueller, S. T. (2012). Executive function on the psychology experiment building language tests. *Behavior Research Methods*, *44*(1), 110–123. https://doi.org/10.3758/s13428-011-0096-6

Piper, B. J., Mueller, S. T., Geerken, A. R., Dixon, K. L., Kroliczak, G., Olsen, R. H. J., & Miller, J. K. (2015). Reliability and validity of neurobehavioral function on the psychology experimental building language test battery in young adults. *PeerJ*, *3*, e1460. https://doi.org/10.7717/peerj.1460

Poehner, M. E., & Lantolf, J. P. (2013). Bringing the ZPD into the equation: Capturing L2 development during computerized dynamic assessment (C-DA). *Language Teaching Research*, *17*, 323–342. https://doi.org/10.1177/1362168813482935

Posner, M. I., Sheese, B. E., Odludaş, Y., & Tang, Y. (2006). Analyzing and shaping human attentional networks. *Neural Networks*, *19*, 1422–1429. https://doi.org/10.1016/j.neunet.2006.08.004

Pureza, J. R., Gonçalves, H. A., Branco, L., Grassi-Oliveira, R., & Fonseca, R. P. (2013). Executive functions in late childhood: Age differences among groups. *Psychology & Neuroscience*, *6*, 79–88. https://doi.org/10.3922/j.psns.2013.1.12

Raven, J., Raven, J. C., & Court, J. H. (2000). *Standard progressive matrices*. Oxford, England: Psychology Press.

Reitan, R. M. (1955). The relation of the trail making test to organic brain damage. *Journal of Consulting Psychology*, *19*, 393–394. https://doi.org/10.1037/h0044509

Resing, W. C. M. (1993). Measuring inductive reasoning skills: The construction of a learning potential test. In J. H. M. Hamers, K. Sijtsma, & A. J. J. M. Ruijssenaars (Eds.), *Learning potential assessment: Theoretical, methodological and practical issues* (pp. 219–241). Amsterdam, the Netherlands: Swets & Zeitlinger Inc.

Resing, W. C. M. (2000). Assessing the learning potential for inductive reasoning in young children. In C. S. Lidz & J. G. Elliott (Eds.), *Dynamic assessment: Prevailing models and applications* (pp. 229–262). Oxford, England: Elsevier.

Resing, W. C. M., & 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*, 579–605. https://doi.org/10.1348/2044-8279.002006

Resing, W. C. M., Elliott, J. G., & Vogelaar, B. (2020). *Assessing potential for learning in school children*. Oxford Research Encyclopedia of Education. https://doi.org/10.1093/acrefore/9780190264093.013.943

Resing, W. C. M., Touw, K. W. J., 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*, 173–191. https://doi.org/10.1080/01443410.2016.1164300

Richland, L. E., & Simms, N. (2015). Analogy, higher order thinking, and education. *Wiley Interdisciplinary Reviews: Cognitive Science*, *6*, 177–192. https://doi.org/10.1002/wcs.1336

Ropovik, I. (2014). Do executive functions predict the ability to learn problem-solving principles? *Intelligence*, *44*, 64–74. https://doi.org/10.1016/j.intell.2014.03.002

Sternberg, R. J. (1985). *Beyond IQ: A triarchic theory of human intelligence*. New York, NY: Cambridge University Press.

Sternberg, R. J., & Grigorenko, E. L. (2002). *Dynamic testing*. New York, NY: Cambridge University Press.

Stevenson, C. E., Heiser, W. J., & Resing, W. C. M. (2013). Working memory as a moderator of training and transfer of analogical reasoning in children. *Contemporary Educational Psychology*, *38*, 159–169. https://doi.org/10.1016/j.cedpsych.2013.02.001

Stevenson, C. E., Touw, K. W. J., & Resing, W. C. M. (2011). Computer or paper analogy puzzles: Does assessment mode influence young children's strategy progression? *Educational and Child Psychology*, *28*, 67–84.

Thibaut, J.-P., & French, R. M. (2016). Analogical reasoning, control and executive functions: A developmental investigation with eye-tracking. *Cognitive Development*, *38*, 10–26. https://doi.org/10.1016/j.cogdev.2015.12.002

Tunteler, E., Pronk, C. M. E., & Resing, W. C. M. (2008). Inter- and intra-individual variability in the process of change in the use of analogical strategies to solve geometric tasks in children: A microgenetic analysis. *Learning and Individual Differences*, *18*, 44–60. https://doi.org/10.1016/j.lindif.2007.07.007

Tzuriel, D., & Galinka, E. (2000). *The conceptual and perceptual analogical modifiability (CCPAM) test: Closed analogies-instruction manual*. Ramat-Gan, Israel: School of Education, Bar-Ilan University.

Tzuriel, D., & Shamir, A. (2002). The effects of mediation in computer assisted dynamic assessment. *Journal of Computer Assisted Learning*, *18*, 21–32. https://doi.org/10.1046/j.0266-4909.2001.00204.x

Van der Sluis, S., De Jong, P. F., & Van der Leij, A. (2007). Executive functioning in children, and its relations with reasoning, reading, and arithmetic. *Intelligence*, *35*, 427–449. https://doi.org/10.1016/j.intell.2006.09.001

Veerbeek, J., Vogelaar, B., Verhaegh, J., & Resing, W. C. M. (2019). Process assessment in dynamic testing using electronic tangibles. *Journal of Computer Assisted Learning*, *35*, 127–142. https://doi.org/10.1111/jcal.12318

Viterbori, P., Usai, M. C., Traverso, L., & De Franchis, V. (2015). How preschool executive functioning predicts several aspects of math achievement in grades 1 and 3: A longitudinal study. *Journal of Experimental Child Psychology*, *140*, 38–55. https://doi.org/10.1016/j.jecp.2015.06.014

Vogelaar, B., Bakker, M., Hoogeveen, L., & Resing, W. C. M. (2017). Dynamic testing of gifted and average-ability children's analogy problem solving: Does executive functioning play a role? *Psychology in the Schools*, *54*, 837–851. https://doi.org/10.1002/pits.22032

Vogelaar, B., Resing, W. C. M., Stad, F. E., & Sweijen, S. W. (2019). Is planning related to dynamic testing outcomes? Investigating the potential for learning of gifted and average-ability children. *Acta Psychologica*, *196*, 87–95. https://doi.org/10.1016/j.actpsy.2019.04.004

Vogelaar, B., Sweijen, S. W., & Resing, W. C. M. (2019). Gifted and average-ability children's potential for solving analogy items. *Journal of Intelligence*, *7*, 19. https://doi.org/10.3390/jintelligence7030019

Weatherholt, T. N., Harris, R. C., Burns, B. M., & Clement, C. (2006). Analysis of attention and analogical reasoning in children of poverty. *Journal of Applied Developmental Psychology*, *27*, 125–135. https://doi.org/10.1016/j.appdev.2005.12.010

Wu, H. M., Kuo, B. C., & Wang, S. C. (2017). Computerized dynamic adaptive tests with immediately individualized feedback for primary school mathematics learning. *Educational Technology & Society*, *20*, 61–72.

Zook, N. A., Davalos, D. B., DeLosh, E. L., & Davis, H. P. (2004). Working memory, inhibition, and fluid intelligence as predictors of performance on tower of Hanoi and London tasks. *Brain and Cognition*, *56*, 286–292. https://doi.org/10.1016/j.bandc.2004.07.003

How to cite this article: Vogelaar B, Veerbeek J, Splinter SE, Resing WCM. Computerized dynamic testing of children's potential for reasoning by analogy: The role of executive functioning. *J Comput Assist Learn*. 2021;37:632–644. https://doi.org/10.1111/jcal.12512

APPENDIX A.

p-values and item-total correlations of the pre- and post-test

	p-values accuracy			*p*-values transformations			Item-total correlations		
		Post-test			Post-test			Post-test	
Item	Pre-test	Dynamic testing	Control	Pre-test	Dynamic testing	Control	Pre-test	Dynamic testing	Control
1	.34	.72	.50	.49	.73	.60	.59	.63	.85
2	.58	.75	.56	.67	.81	.63	.70	.76	.95
3	.22	.81	.59	.53	.85	.65	.54	.89	.96
4	.14	.78	.59	.44	.82	.64	.53	.76	.96
5	.44	.59	.25	.71	.84	.66	.66	.42	.56
6	.48	.78	.56	.64	.87	.66	.74	.82	.90
7	.25	.75	.56	.54	.84	.66	.77	.80	.95
8	.27	.63	.44	.62	.80	.60	.64	.76	.82
9	.28	.75	.62	.60	.85	.67	.60	.86	.92
10	.19	.81	.47	.58	.86	.66	.54	.86	.82
11	.09	.59	.59	.61	.81	.67	.37	.68	.87
12	.25	.72	.41	.58	.81	.65	.75	.83	.77
13	.25	.34	.38	.57	.73	.56	.73	.59	.77
14	.20	.66	.53	.53	.78	.65	.69	.81	.88
15	.27	.34	.13	.57	.68	.53	.69	.56	.47