Dynamic testing and excellence: unfolding potential
Vogelaar, B.

Citation

Vogelaar, B. (2017, January 18). *Dynamic testing and excellence: unfolding potential*. Retrieved from https://hdl.handle.net/1887/45569

Cover Page

The handle http://hdl.handle.net/1887/45569 holds various files of this Leiden University dissertation.
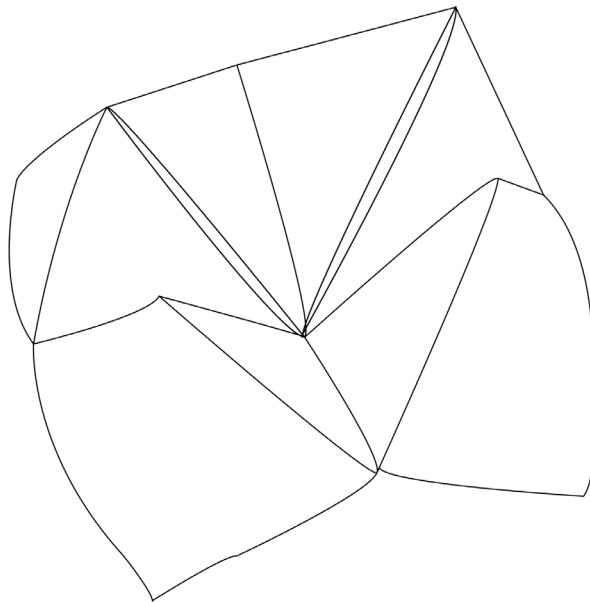
Author: Vogelaar, B.
Title: Dynamic testing and excellence: unfolding potential
Issue Date: 2017-01-18

CHAPTER 5

Changes over time and transfer of analogy-problem solving of gifted and non-gifted children in a dynamic testing setting

Bart Vogelaar
Wilma C. M. Resing

Abstract

This study examined differences in transfer of analogical reasoning after analogy problem-solving between 40 gifted and 95 average-ability children (aged 9-10 years old), utilising dynamic testing principles. This approach was used in order to examine potential differences between gifted and average-ability children in relation to progression after training, and with regard to the question whether training children in analogy problem-solving elicits transfer of analogical reasoning skills to an analogy construction-task. Children were allocated to one of two experimental conditions: either children received unguided practice in analogy problem-solving, or they were provided with this in addition to training incorporating graduated prompting techniques. The results showed that gifted and average-ability children who were trained made more progress in analogy problem-solving than their peers who received unguided practice experiences only. Gifted and average-ability children were found to show similar progression in analogy problem-solving, and gifted children did not appear to have an advantage in the analogy-construction transfer task. The dynamic training seemed to bring about no additional improvement on the transfer task over that of unguided practice experiences only.

5.1. Introduction

Applying knowledge to a new context is an important necessity in order for gained knowledge and skills to be of use in everyday life outside the classroom context (Day & Goldstone, 2012), and is therefore one of the main aims of education. Groups of children have been found to differ in the extent to which they transfer learned knowledge and skills. Lower income students have, for example, been found to have more difficulty transferring knowledge and skills than their middle income peers (Alexander, Entwisle, & Olson, 2001). Other groups of learners have, in contrast, been found to have an advantage in transfer of learning, one such group being the gifted. Gifted children have long been thought to have an excellent ability to transfer learning to a new situation (e.g., Renzulli, Smith, White, Callagan, Hartman et al., 1997).

Considering the importance of transfer, it comes as no surprise that it has been studied for more than a hundred years (Engle, 2012). However, various studies have shown that eliciting transfer of learning to new contexts proves difficult (Bransford, Brown, & Cocking, 2001; Day & Goldstone, 2012; Gick & Holyoak, 1983), possibly due to its complex (Nokes-Malach & Mestre, 2013), and idiosyncratic nature (Kyllonen, Lohman, & Snow, 1984). Transfer has been noted for its potential to reveal important insights into children's potential for learning (Bosma & Resing, 2006; Ferrara, Brown & Campione, 1986). Therefore, in dynamic testing, transfer of newly acquired knowledge and skills is one of the measures used to gain insight into a child's potential for learning (Campione & Brown, 1987; Sternberg & Grigorenko, 2002; Tzuriel, 2007). In contrast with conventional static testing, dynamic testing is a form of testing that incorporates feedback and instruction, sometimes tailored to the individual, into the testing process (Elliott, 2003; Jeltova, Birney, Fredine, Jarvin, Sternberg et al., 2007), and is focused on the potential for learning, rather than on previously acquired skills and knowledge (Resing & Elliott, 2011).

In the present study, we applied dynamic testing principles to examine whether cognitively gifted and average-ability 9 and 10 year old children would show differential changes in analogy problem-solving, and differential patterns in their ability to transfer analogy problem-solving skills to an analogy-construction task. All the children in the present study received opportunities for unguided practice in analogy problem-solving. Half of the children, however, received an additional training in analogy problem-solving, which enabled us to investigate whether training would lead to more changes over time in problem-solving than unguided practice, and facilitate transfer of the learned skills.

Dynamic testing outcomes are assumed to provide a more detailed picture

of a child's cognitive potential (Elliott, Grigorenko & Resing, 2010), strengths and weaknesses (Jeltova et al., 2007), than conventional, static testing procedures, such as intelligence or school aptitude tests (Elliott, 2003). This form of testing has been found to be especially beneficial for special populations, such as ethnic minority, or learning disabled children (Robinson-Zañartu & Carlson, 2013; Sternberg & Grigorenko, 2002). Perhaps due to these reasons, the large majority of research into dynamic testing has focused on special groups of children. Studies into dynamic testing of gifted children, however, are few. Findings of such studies have revealed that gifted children not only outperform their non-gifted peers, but also showed significantly more improvement (Calero, García-Martin, & Robles, 2011). Moreover, young gifted children were found to have a more extensive zone of proximal development, to learn new skills faster, and to be better at generalising new knowledge (Kanevsky, 1990, 2000).

Dynamic tests often employ inductive reasoning tasks (e.g., Ferrara et al., 1986; Resing, 2000; Resing & Elliott, 2011). Inductive reasoning is assumed to be related to a large variety of higher-order cognitive skills and processes (Csapó, 1997), including general intelligence (Klauer & Phye, 2008), problem solving (Richland & Burchinal, 2012), and applying knowledge and skills (Goswami, 2012). Analogical reasoning, a subtype of inductive reasoning, is considered to play a central role in cognitive development (Klauer & Phye, 2008; Pellegrino & Glaser, 1982), and develops significantly throughout childhood (e.g., Leech, Mareschal, & Cooper, 2008). Moreover, children's analogical reasoning ability can be characterised by large individual differences (e.g. Siegler & Svetina, 2002). Not only do older children perform better than younger children (Csapó, 1997), children with strong cognitive capacities, such as gifted and talented children, are also found to achieve higher scores on analogical reasoning tasks (e.g., Caropreso & White, 1994).

Several studies in the field of dynamic testing have revealed that training incorporating graduated prompting techniques, can lead to improvement in reasoning by analogy (Bosma & Resing, 2006; Ferrara et al., 1986; Stevenson, Heiser, & Resing, 2013). Graduated prompting techniques, as used in the present study, refer to a form of an intervention in which children are provided with prompts each time they make a mistake in problem solving. In the current study, prompts are tailored to each individual problem to be solved, and become more specific gradually, from metacognitive to cognitive prompts and modelling (Resing, 2000; Resing & Elliott, 2011). Graduated prompting techniques are used increasingly in combination with a pretest-training-posttest format, as in the present study, a specific form of dynamic testing that allows for structured measuring of children's

progression in learning (e.g., Ferrara et al., 1986; Resing & Elliott, 2011; Resing, Bakker, Pronk, & Elliott, 2016).

Analogical reasoning involves defining and deciding that two problem-solving situations are similar, and, ultimately, successfully transferring previously problem-solving experiences to new situations that can be, partially, dissimilar. Unsurprisingly, reasoning by analogy is considered to be closely related to the ability to transfer (Gentner, Holyoak, & Kokinov, 2001); both require that one observes an analogy or similarity between two problems (Chi & VanLehn, 2012; Holyoak, 1984). In general, two factors have been proposed to play a role in transfer: the content, the exact problem that is being transferred (Barnett & Ceci, 2002, and the context (Klahr & Chen, 2011), which refers to the different domains from and to which the problem is being transferred. Researchers often distinguish in different types of transfer on the basis of the surface similarity of the base and target problem, including near versus far transfer (Mestre, 2005), and surface versus deep transfer (Forbus, Gentner, & Law, 1995). If the base and target share few surface similarities, and are thus less similar, the more cognitively demanding the process of transfer becomes (Gentner, Loewenstein, & Hung, 2007). Transferring effectively involves mastery of the task to be transferred (Siegler, 2006), and a deep, rather than surface understanding of the task at hand is required for deep transfer (Barnett & Ceci, 2002).

Several studies have shown that great variability exists in the extent to which children can transfer knowledge to new domains (e.g., Tunteler & Resing, 2010). It is often assumed that gifted children have a cognitive advantage, which enables them to transfer knowledge more efficiently than their non-gifted peers (e.g., Klavir & Gorodetsky, 2001; Zook & Maier, 1994). Research into the transfer ability of this group of children has revealed that on near transfer tasks gifted children's performance seems similar to their non-gifted peers (Carr, Alexander, & Schwanenflugel, 1996). In far transfer tasks however, gifted children were found to outperform their non-gifted peers (Geake, 2008; Kanevsky, 2000). Kanevsky (1990) reported that, after learning new strategies, gifted learners spontaneously transferred these strategies to new learning contexts. The underlying processes facilitating transfer in the gifted population are not yet fully understood, but Carr and colleagues (1996) suggest that gifted learners are more likely to transfer their acquired strategies to other domains, as they show an elaborate understanding, and make more use of complex strategies.

As eliciting transfer is challenging (e.g., Barnett & Ceci, 2002), a variety of studies have been carried out investigating whether training facilitating deep understanding or mastery of a task could promote transfer. Several studies have

5

revealed that training children in solving inductive reasoning problems led to higher levels of generalising skills learned during training to similar and dissimilar problems in the same inductive reasoning domain (e.g., Harpaz-Itay, Kaniel, & Ben-Amram, 2006; Roth-van der Werf, Resing, & Slenders, 2002; Tzuriel, 2007). In the present study, we utilised a 'reversal' procedure to measure transfer. Having had practice opportunities, or practice opportunities in combination with a short training in analogy problem-solving, participants were asked to construct their own analogy items, similar to the ones they had solved before, which then had to be solved by the examiner (Bosma & Resing, 2006; Kohnstamm, 2014; Stevenson et al., 2013). As such, this task required a reversal of roles.

In order to promote transfer of problem-solving strategies practiced or trained, we kept the surface features of our analogy construction task similar to those of the open-ended visuo-spatial geometrical analogy items children solved before (Resing et al., 2016), assuming that children would use previously acquired knowledge and skills in their constructed analogies (Day & Goldstone, 2012). Previous research, however, has shown that despite these similarities in surface structure, the analogy construction task is a challenging and difficult task for children (Bosma & Resing, 2006; Tzuriel & George, 2009).

The present study had two main aims. Although consideration of the occurrence of transfer was our primary research aim, we were also interested in whether children's analogy problem-solving would improve differentially. Firstly, we sought to examine children's (differential) potential for learning. We expected that training by dynamic testing would lead to more change in children's analogy problem-solving than unguided practice only. We anticipated larger progression in accuracy scores of the children who were dynamically trained than the children wo received unguided practice only (Stevenson et al., 2013; Tunteler, Pronk, & Resing, 2008). We further anticipated that progression in accuracy would be larger for gifted than average-ability children, and there would be a significant interaction between session, condition and ability group for the accuracy scores (Calero et al., 2011).

We also considered the time it took children to complete all of the items of a test session. We expected that results for children in the unguided practice would show a decrease in completion time, but not for those in the dynamic testing condition, as we expected that training would lead children to spend time on strategic considerations (Resing, Tunteler, & Elliott, 2015). We also expected a significant three-way interaction of session x condition x ability group as to children's completion time, and hypothesised that gifted children would be more time efficient than their average-ability peers, considering they are assumed to

be better in self-regulation (Calero, García-Martín, Jiménez, Kazén, & Araque, 2007).

Our second research question concerned the transfer of learned skills. As proposed by Clerc, Miller, and Cosnefroy (2014), in-depth assessment of transfer requires the measuring of both performance, as well as the degree of transfer achieved. Therefore, in the current study, we focused on both the transfer accuracy scores, as well as on the difficulty level of the analogies constructed by the children. We expected that, in comparison with children in the unguided practice condition, trained children would show higher levels of transfer accuracy scores; as well as difficulty levels of accurately constructed analogies (Resing, 1997; Roth-van der Werf et al., 2002). In addition, we expected that gifted children would show a higher degree of transfer than their average-ability peers (Geake, 2008; Kanevsky, 1990). We further explored whether children's transfer performance and degree of transfer could be predicted by their analogy problem-solving accuracy scores (Alexander & Murphy, 1999).

5.2. Method

Participants

In the present study, 135 children participated, 62 boys and 73 girls, ranging in age from 9 years and 3 months to 10 years and 11 months (M=10;10; SD=0;6). All the participants were born in the Netherlands, and attended either a mainstream primary school, or a special setting for gifted and talented children in the western part in the Netherlands. All schools participated on a voluntary basis. Gifted children were over-sampled and identified on the basis of a qualitative judgment of parents and teachers regarding their giftedness. Additionally, all of the children in our gifted sample each scored at, or above the 90th percentile on the Raven's Progressive Matrices Test (Raven, 1981). Written permission of parents and school was obtained for each child prior to participation in the current study. Nine children dropped out in the course of the study, as they did not participate in each test session.

Design

The study used a two-session (pre-test, post-test) repeated measures randomised blocking design with two treatment conditions: dynamic testing versus unguided practice (see Table 1). The children in the dynamic testing condition received two short training sessions between pre-test and post-test, whereas the children in the unguided practice condition did not receive any practice or training opportunities. Before the pre-test, the Raven Progressive Matrices Test (Raven, 1981) was administered to allocate children to the two treatment conditions. Only the children who had obtained a Raven percentile

score of at least 90, were included in the "gifted" condition, the other children in the "average-ability" condition.

Raven scores were used to ensure that any differences in initial reasoning ability were as small as possible across the children in the dynamic testing and unguided practice conditions. Within the two ability groups, pairs of children with equal scores (blocking) were randomly assigned to the dynamic testing or unguided practice condition, resulting in four subgroups of children: gifted dynamic testing (n=22), gifted unguided practice (n=18), average-ability dynamic testing (n=47) and average-ability unguided practice (n=48).

Table 1. Overview of the design

Condition	Groups		Dynamic/static test				
		Raven	Pre-test	Training 1	Training 2	Post-test	Transfer
Dynamic testing	Gifted (n=22)	X	X	Dynamic training 1	Dynamic training 2	X	X
	Average-ability (n=47)						
Unguided practice	Gifted (n=18)	X	X	---	---	X	X
	Average-ability (n=48)						

Materials

Raven Progressive Matrices Test. The Raven Progressive Matrices Test (Raven, 1981), a non-verbal intelligence test measuring fluid intelligence, was used as a blocking instrument. The Raven test results were shown to have a high level of internal consistency in several studies as shown by split-half-coefficients of r =.91 (Raven, 1981).

Dynamic test of analogical reasoning.

Pre-test and post-test. The dynamic test utilised visuo-spatial geometric analogies of the type A:B::C:?? of varying difficulty, part of a test battery developed by Hosenfeld, Van den Boom, & Resing (1997), and adapted for further use by Tunteler et al., (2008). Six basic geometrical shapes were used in the construction of the analogies: squares, triangles, hexagons, pentagons, circles, and ellipses (see Figure 1 for an example of a difficult analogy item). The original analogy test items were constructed by a maximum of five possible transformations: changing position, adding or subtracting an element, changing size, halving, and doubling.

As the original item-sets have been used for young children of various ages, but not for children from the age of nine, the items used in the current study were adapted by adding extra transformations, including rotation and colour. The test was administered as an open-ended paper-and-pencil test, and children were asked to draw their answers.
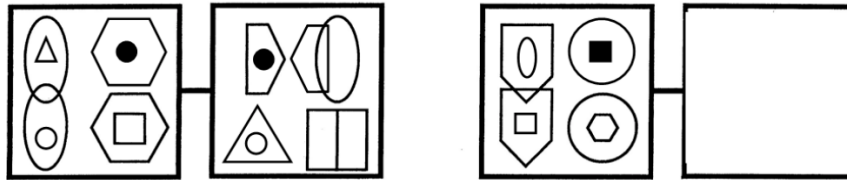

Figure 1. Example of a difficult analogy item

Both the pre-test and the post-test consisted of 21 items of varying difficulty. For pre- and post-test, parallel versions were constructed by keeping the difficulty levels of the items the same, as well as the order in which the items of varying difficulty were presented. Participants did not receive any feedback on or help with their given answers during the pre and post-test, but received minimal instructions that only specified the children had to solve puzzles by filling the empty square with the appropriate shapes. In our sample of participants, the pre-test was found to have high internal consistency (Cronbach's α=.85) for the accuracy scores.

Dynamic training. Two short training sessions each consisting of 6 new analogy items were administered between the pre-test and post-test to participants in the dynamic testing condition. The training sessions employed graduated prompting techniques used in earlier studies (e.g., Campione & Brown, 1987; Resing, 2000; Resing & Elliott, 2011). These involve the provision of a number of prompts when the child makes an error in problem-solving. All prompts were administered hierarchically: starting with four very general metacognitive prompts, followed by four specific cognitive prompts, tailor-made for each item. As each new prompt became progressively more specific, this procedure enabled measurement of the child's need for differing degrees of help in order to solve the problem presented. Both training sessions consisted of eight prompts in total, which were only administered after indication that a child could not solve the problem independently.

After each prompt, children were asked to draw the solution of the analogy, and check whether their solution was correct. If a child had not solved the analogy after the seventh prompt had been administered, the examiner modelled the correct answer. After responding, participants were asked to explain why they thought their answer was correct. Finally, the tester provided a correct self-explanation. Figure 2 shows a flowchart of the training procedure.
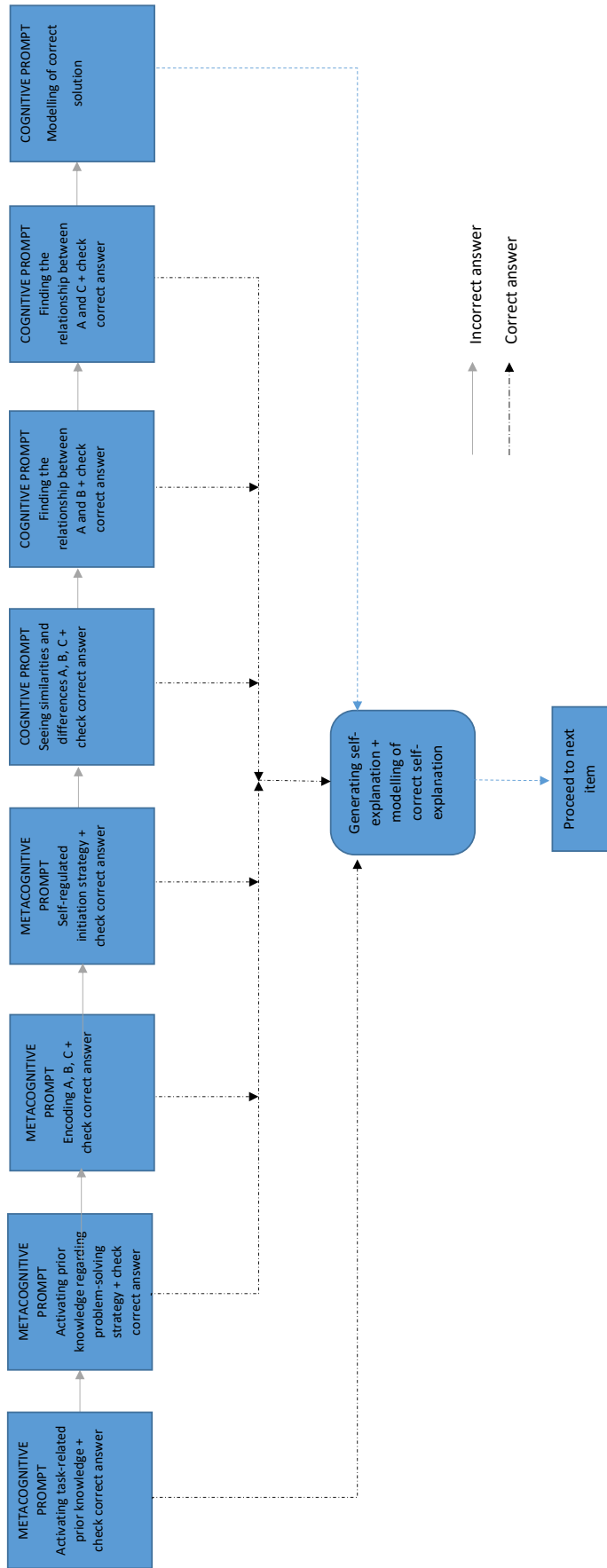
Figure 2. Flowchart of the graduated prompts training protocol.

Transfer: Analogy construction task. The ability to transfer previously practised or learned analogy problem-solving was measured by means of an analogy construction task. As part of this task, children were asked to construct their own analogy, so the examiner could solve it. In a sense, the roles were reversed, and the child became the teacher (Bosma & Resing, 2006). Participants were provided with four squares, similar to those used in the previous test sessions, but then empty, and instructed that they could utilise any of the geometric shapes they had seen in prior sessions, and to provide instructions to the examiner on how to solve the analogy. Deeper understanding of analogical reasoning principles is required to be able to construct a correct analogy (Harpaz-Itay et al., 2006). Children were asked to complete two reversal trials, and thus construct two analogies. For both tasks, the children were given short, general instructions only to enhance spontaneous problem-solving behaviour. After construction of the analogy item, the child had to ask the examiner to solve the item, and, on completion of the analogy by the examiner, then had to explain why this was the correct answer.

General procedure

Children were tested once weekly, in accordance with the schools' availability, over a period of six consecutive weeks. All parts of the present study were administered individually, following standard, protocolled instruction. At the beginning of the pre-test, training and post-test sessions, the children were provided with the six geometrical shapes used in the analogies, and, in cooperation with the examiner named each shape, after which the examiner asked the child to draw the shapes below the printed shapes, staying as close to the original as possible (Tunteler et al., 2008). It was assumed that this procedure helped activate children's prior knowledge, ensured that the test leader and child used the same terminology when addressing the geometric shapes, and, in doing so, facilitated the scoring procedure.

The solutions of the analogy items that the children had drawn during pretest, training and post-test were collected, and completion times of the pre-test and post-test were recorded. The analogies the children had constructed as part of the transfer task were collected, scored on accuracy as well as on the number of transformations the item consisted of.

Scoring and analyses

The outcome variables of the pre-test and post-test sessions consisted of the accuracy score, the total number of correct items per session, and the completion time, the time (in seconds) it took each child to solve all the items of the pre-test and post-test.

We examined two outcome variables of the analogy construction task: transfer accuracy, and difficulty level. The first outcome variable was the sum of accurately constructed analogies (range 0-2). The second outcome measure was the transfer difficulty, calculated by means of the equation correctness of the analogy constructed (1/0) x the number of transformations used in the construction of the analogy (1-8; Mulholland, Pellegrino, & Glaser, 1980; Stevenson et al., 2013). The total difficulty level score (range 0-14) was calculated by adding the difficulty level scores of both items. Both outcome variables were considered to be ordinal, violating the assumptions of least-squares regression. Therefore, we conducted ordinal logistic regression analysis (Agresti, 2010). The regression analyses included the following predicting variables: condition, ability group, condition x ability group, post-test accuracy score, and condition x post-test accuracy score.

5.3. Results

Initial group comparisons

Prior to analysing our research questions, we evaluated possible differences between the two experimental conditions, and ability groups, respectively. The children in the two conditions did not differ in their age, or initial reasoning performance (Raven accuracy score). Children in the gifted and non-gifted groups also did not differ in age, but did in their initial reasoning performance ($p<.001$). We further evaluated possible differences in pre-test performance, and found no significant differences in accuracy scores, or in completion time between children in the two experimental conditions. Gifted and non-gifted children were found to differ on their accuracy scores ($p<.001$), but not on their completion time. Basic statistics for the measures used in the current study are provided in Table 2.

Table 2. Basic statistics of the analogical measurements, divided by condition and ability group

			Dynamic testing			Unguided practice		
			Gifted	Average-ability	Total	Gifted	Average-ability	Total
		N	22	47		18	48	
Raven	Accuracy	M	49.73	38.94	42.38	49.89	38.69	41.74
	scores	SD	2.51	5.90	7.15	2.78	6.18	7.41
Pre-test	Accuracy	M	10.77	5.91	7.46	9.83	7.08	7.83
	scores	SD	4.38	4.14	4.77	4.29	3.71	4.04
	Completion	M	1267.45	1207.52	1266.63	1331.09	1199.09	1235.09
	time	SD	302.80	564.37	494.54	367.80	471.58	446.87
Post-test	Accuracy	M	15.41	11.83	12.97	11.06	8.27	9.03
	scores	SD	3.42	4.89	4.76	5.13	4.15	4.57
	Solving-time	M	1055.05	1104.34	1088.62	1075.78	1027.42	1040.61
		SD	226.26	329.50	299.65	270.99	325.55	310.34
Transfer	Accuracy	M	1.77	1.66	1.62	1.94	1.50	1.70
	scores	SD	.43	.60	.65	.24	.72	.55
	Complexity	M	8.00	7.40	7.59	6.89	5.46	5.85
		SD	3.10	3.31	3.23	2.70	3.36	3.24

5

Analogy problem-solving

Our first research question concerned changes in analogy problem-solving. Changes over time were examined by means of two repeated measures ANOVAs with one within-subjects factor Session (Session 1-2), and two between-subjects factors Condition (dynamic testing versus
unguided practice) and Ability group (gifted versus average-ability). Children's accuracy scores and completion time at Sessions 1 and 2 were used as the dependent variables. In Table 3, the main and interaction effects of the repeated measures ANOVAs are provided. Results for accuracy scores revealed significant main effects of Session ($p<.0001$, $\eta_p^2=.39$), and, most important for answering our hypothesis, a significant Session x Condition ($p<.0001$, $\eta_p^2=.20$) interaction, but no significant Session x Ability group ($p=.39$), or Session x Condition x Ability group interaction ($p=.36$). Inspection of Table 3 led to the conclusion that, only partially in accordance with our hypotheses, dynamically tested children, irrespective of their ability group, showed significantly greater progression in solving analogies than control-group children. The slopes of the progression lines of the two dynamically tested groups of children did not significantly differ, indicating that children in both ability groups made comparable progress in accuracy although they started at different levels. The between-subjects effects of Ability group for

accuracy supported this finding, $F(1,131)=23.42$, $p<.0001$, $\eta_p^2=.15$.

Table 3. Results of the repeated measures ANOVAs for the accuracy scores, and completion time

	Wilks' λ	F	p	η_p^2
Accuracy scores				
Session	.62	82.09	<.0001	.39
Session x Condition	.80	32.39	<.0001	.20
Session x Ability group	.99	.76	.39	.01
Session x Condition x Ability group	.99	.84	.36	.01
Completion time				
Session	.86	21.75	<.0001	.14
Session x Condition	1.00	.49	.49	.00
Session x Ability group	.99	1.47	.23	.01
Session x Condition x Ability group	1.00	.03	.87	.00

A second aspect of children's analogy solving concerned the time they needed to complete all of the tasks of a test session. Although we expected that completion time would decrease for the children in the unguided practice condition, but not for the trained children, the repeated measures ANOVA showed only a significant main effect of session ($p<.0001$, $\eta_p^2=.14$), but no significant interaction effects (see Table 3). Contrary to our expectations, all groups of children showed a comparable decrease in their completion-time from pre-test to post-test.

These results led us to conclude that training leads to more improvement in accuracy than practice opportunities, as assumed, but, unexpectedly, that training and unguided practice both led to a decrease in solving time. Gifted and average-ability children seemed to differ in terms of the number of items solved correctly, as expected, with an advantage for those who were gifted, but, in contrast to our hypotheses, not in terms of change from pre-test to post-test, and training benefits in relation to both accuracy scores and completion time. Of course, individual differences in scores and changes in scores are large for children in both ability groups and conditions.

Transfer of analogy problem-solving

Our second research question concerned children's performance on the analogy construction transfer task. Ten children were unable to construct any accurate analogies, with eight children constructing items that were partial analogies, and two constructing items that were non-analogical. Out of 135

children, 27 and 98 children, accurately constructed either one, or two analogies, respectively (see Table 4). As a further exploration of the data, the children who had been able to construct accurate analogies were divided in three groups, based on the difficulty level of the analogies they constructed: low, medium, and high. Chi square tests revealed that children in the unguided practice condition constructed more analogies of a low difficulty level than the trained children (χ^2(1)=4.57, p=.03), that trained children designed more difficult items than non-trained children (χ^2(1)=5.49, p=.02), while the children who had constructed analogies of medium difficulty level were distributed evenly (χ^2(1)=.53, p=.47). These findings revealed a first indication of the effect of training on transfer accuracy and difficulty level.

In addition, two Mann-Whitney U tests were conducted, one for each condition, to explore whether gifted and average-ability children were distributed evenly in relation to the difficulty level of the analogies they constructed. The results revealed that gifted and average-ability children were distributed evenly across the three difficulty level groups in the dynamic testing (U=448, z=-.53, p=.60), and the unguided practice condition (U=392, z=.24, p=.81).

5

Table 4. Children's transfer accuracy and difficulty level, divided by condition and ability group

	Dynamic testing			Unguided practice		
	Gifted	Average-ability	**Total**	Gifted	Average-ability	**Total**
Inaccurate analogies						
Non-analogical	0	0	**0**	0	2	**2**
Partial analogical	0	3	**3**	0	5	**5**
Accurate analogies						
Low difficulty (2-6 transformations)	4	7	**11**	8	18	**26**
Medium difficulty (7-9 transformations)	10	16	**26**	5	16	**21**
High difficulty (10-16 transformations)	8	20	**28**	5	8	**13**

An ordinal regression analysis was conducted to examine whether the number of accurately constructed analogies could be predicted by condition (dynamic testing versus unguided practice), ability group (gifted versus average-ability), and post-test accuracy. The results (see Table 5) revealed that the post-test accuracy score (p=.001) could significantly predict transfer performance. Although neither condition (p=.18), nor ability group (p=.50) significantly

contributed to prediction, the condition x ability group interaction did (*p*=.04). The interaction was, contrary to our hypotheses, in the wrong direction.

A second ordinal regression analysis was conducted to examine whether transfer difficulty level could be predicted by condition, ability group and post-test accuracy. The findings indicated that only the post-test accuracy scores contributed significantly to prediction (*p*=.001). In sum, children's post-test accuracy score seemed to be a good predictor of transfer accuracy and difficulty level, regardless of training or ability group. Unexpectedly, however, the gifted children who received unguided practice seemed to outperform the gifted children who were trained in terms of transfer accuracy. An exploration of the quality of the constructed analogies suggested that differences between children were found mainly in the items of lower and higher difficulty group, with training seemingly facilitating construction of more difficult items.

Table 5. Results of the regression analyses for transfer accuracy and difficulty level (correct x transformations)

	b (SE)	Exp *(β)*	Exp *(β)*	*χ²*	*P*
Accurate analogies [a]					
Condition	2.16 (1.59)	8.64	8.64	1.83	.18
Ability group	.48 (.71)	1.61	1.61	.45	.50
Condition x Ability group	-2.75 (1.33)	.06	.06	4.29	.04
Post-test accuracy score	.22 (.06)	1.25	1.25	11.82	.001
Condition x Post-test accuracy score	.09 (.10)	1.09	1.09	.73	.39
Transfer difficulty [a]					
Condition	.14 (1.07)	1.15	1.15	.02	.90
Ability group	.30 (.48)	1.35	1.35	.39	.53
Condition x Ability group	-.69 (70)	.50	.50	.98	.32
Post-test accuracy score	.16 (.05)	1.18	1.18	10.91	.001
Condition x Post-test accuracy score	-.01 (.07)	1.00	1.00	.01	.94

Note. [a] Ordinal logistic regression.

5.4. Discussion

The focus of the present study was two-fold. We examined gifted and average-ability children's progression in analogy problem-solving after dynamic training or unguided practice. We also focused on whether a dynamic training would facilitate children's performance and degree of transfer, and whether gifted and average-ability children show differences in their transfer accuracy and

difficulty level.

We first looked into children's potential for learning. The children who, in addition to unguided practice experiences, also received dynamic training showed steeper progression in accuracy than the children who were not trained, indicating that testing children's ability dynamically shows a more complete picture of their cognitive potential than testing statically (see e.g., Robinson-Zañartu & Carlson, 2013). We also focused on potential differences between gifted and average-ability children in relation to changes in analogy problem-solving performance. Our findings suggest, as expected (e.g., Calero et al., 2011), that gifted children outperformed their average-ability peers in terms of accuracy in analogy problem-solving. They did not, however, show differential progression in accuracy or reduction in completion time, which is in contrast to earlier findings (e.g., Kanevsky, 1990, 2000).

In addition, children who were trained showed similar levels of reduction in time needed to solve the analogies to their peers who did not receive training. Although training seemed to lead to more advanced analogy problem-solving it did not lead to children spending more time on completing the tasks. In earlier research (Resing et al., 2015), it was posited that this might be due to children devoting more time to strategic considerations as a consequence of training. In the current study, however, children had to draw their own answers, which required substantial time and usage of motor skills. Completion time seemed to be dependent on other factors, such as children's fine motor skills. The reduction in completion time could, therefore, be ascribed to more familiarity with the task and an improvement in fine motor skills needed for the task, rather than other aspects, such as strategic considerations. In future studies, therefore, distinguishing between planning, and task execution time might lead to insights in relation to children's time allocation while solving analogy items, and whether gifted and average-ability children show differential patterns of time allocation and efficiency.

Our second main aim was to explore potential differences in gifted and average-ability children's transfer of practiced or learned skills. We utilised an analogy construction task (Bosma & Resing, 2006; Harpaz-Itay et al., 2006) in order to examine children's transfer accuracy and difficulty level. We expected that the transfer task would be difficult for the children, as it requires deep understanding of the task (e.g., Harpaz-Itay et al., 2006), and that at least some of the children would need training in order to facilitate this deep understanding. First of all, we found that the majority of children could construct an accurate analogy, in contrast with earlier studies in which more children were found to

5

have difficulty with this task (e.g., Resing et al., 2016; Stevenson et al., 2013). The children participating in the current study were slightly older than in these previous studies, so we suggest that this is partly due to developments in their analogical reasoning (Csapó, 1997; Leech et al., 2008). Training, however, could not predict transfer accuracy or transfer difficulty level.

Further, children were divided in groups on the basis of the degree to which they could transfer, looking more closely at the difficulty level of the constructed analogy items. We found that the group of children that had constructed analogy items with a high difficulty level contained significantly more children who were dynamically trained, while the group that had constructed low difficulty analogy items consisted of significantly less dynamically trained children than children who had received unguided practice opportunities. These findings suggest that, for at least some children, training was necessary for them in order to construct the more difficult analogy items. Other children, however, did not need training to construct difficult items, reflecting individual differences between children in relation to their analogy problem-solving and construction skills.

One might assume that the group of children who constructed the more difficult items consisted mostly of the gifted children, but, contrary to our expectations, gifted children were not found to outperform their average-ability peers in transfer accuracy as well as difficulty level. In other dynamic testing studies, gifted children's performance was characterised by significantly more progression in performance (e.g., Calero et al., 2011) as well as higher transfer rates (e.g., Kanevsky, 1990), which led these authors to conclude that these children have a more extensive zone of proximal development. It must be noted that in the current study, children had not been formally identified as gifted by means of full scale IQ testing, but were identified on the basis of their parents' and teachers' judgements as well as their scores on the Raven test. Although the Raven test is considered a reliable measure of general intelligence, perhaps utilising a stricter cut-off score than the 90th percentile used in the current study, or taking into account other factors rather than just cognitive factors (see e.g., Kornilov, Tan, Elliott, Sternberg, & Grigorenko, 2012; Renzulli et al., 1997), would have led to more distinct differences in performance.

Moreover, it must be taken into consideration that the instructions of the transfer task did not specify that children had to construct complex analogies. Instead, instructions were kept to a minimum to elicit spontaneous problem-solving (e.g. Resing et al., 2016). Therefore, it cannot be disregarded that some of the gifted children were not motivated to construct difficult items, due to various reasons. Perhaps, some did not find the task sufficiently challenging to construct

a highly complex item.

Clerc et al. (2014) provide an alternative explanation. These authors stated that self-regulation is strongly associated with the ability to transfer. They postulate that both children with low self-regulation as well as those with high self-regulation, such as gifted children, as for example demonstrated by Calero et al. (2007), can experience difficulty with transfer. According to them, good metacognition might hinder some children's transfer ability, as they do not want to use the strategy they have learned, before having fully mastered the cognitive processes necessary for utilising the strategy. A child's metacognitive knowledge relating to the transferral of a strategy or skill might be ahead of the child's actual ability to apply the strategy or skill. The fact that the gifted children who received unguided practice outperformed, in terms of transfer accuracy, their gifted peers who were trained lends some support to this explanation; perhaps training enhanced these children's metacognitive knowledge, while their actual ability to apply what they had learned in training was not yet at the same level, making these children unwilling to apply the strategies they have learned to a difficult item. Further research, with a larger sample of gifted children, ought to be conducted to further investigate these claims more thoroughly, taking into account, specifically, children's strategy use.

Clerc et al.'s (2014) explanation might also, in part, account for the fact that training, contrary to our expectations, could not predict transfer accuracy or difficulty level. Other explanations could be that the tasks were too difficult for some of the children to achieve deep understanding in a short time-frame (e.g., Tzuriel & George, 2009), or that the training employed in the current study was too short. In future studies, it might be useful to make the training more intensive, for example by increasing the number of sessions, or the number of items per training session (Resing et al., 2016; Tzuriel & George, 2009). More research, however, is necessary to investigate exactly what type of training is beneficial. Considering the individual differences portrayed by children in the current and in previous studies (e.g., Resing et al., 2015, 2016; Stevenson et al., 2013), children might benefit more from training that is more tailored.

Finally, the present study contributed to the existing research into transfer as we investigated both transfer accuracy, and complexity, and, thus, looked into both transfer performance and the degree of transfer obtained by the children (Clerc et al., 2014). These authors postulate that effectiveness of transfer can only be captured by measuring these two aspects. As children's accuracy in analogy problem-solving was found to predict both transfer performance, as well as transfer effectiveness, our findings support the notion that deeper understanding

of analogy problem-solving is required for successful analogy construction (Harpaz-Itay et al., 2006).

The findings from the present study also support Siegler's (2006), and Day and Goldstone's (2012) suggestions that mastery of a skill is a requirement for transfer to occur, especially at the deep level, and that transfer at the deep level is challenging for young children (e.g., Clerc et al., 2014; Resing et al., 2016). Only a number of children could construct more difficult items, the majority of whom had received training in analogy-solving. Children showed considerable individual differences in their progression in accuracy, as well as their performance and effectiveness of transfer, findings that could only partially have been captured by traditional, static testing. In that sense, it seems plausible that dynamic testing might be a valuable instrument in capturing the underlying processes involved in progression in performance, as well as transfer in relation to learned skills.