WILEY

RESEARCH ARTICLE

Dynamic testing of gifted and average-ability children's analogy problem solving: Does executive functioning play a role?

Bart Vogelaar[1] | **Merel Bakker[1]** | **Lianne Hoogeveen[1,2]** |
Wilma C. M. Resing[1] iD

[1]Leiden University

[2]Radboud University Nijmegen

Correspondence
Wilma C. M. Resing, Section Developmental and Educational Psychology, Department of Psychology, Faculty of Social Sciences, Leiden University, P.O. Box 9555, 2300 RB Leiden, The Netherlands.
Email: resing@fsw.leidenuniv.nl

In this study, dynamic testing principles were applied to examine progression of analogy problem solving, the roles that cognitive flexibility and metacognition play in children's progression as well as training benefits, and instructional needs of 7- to 8-year-old gifted and average-ability children. Utilizing a pretest training posttest control group design, participants were split in four subgroups: gifted dynamic testing ($n = 22$), gifted unguided practice ($n = 23$), average-ability dynamic testing ($n = 31$), and average-ability unguided practice ($n = 37$). Results revealed that dynamic testing led to more advanced progression than unguided practice, and that gifted and average-ability children showed equivalent progression lines and instructional needs. For children in both ability categories, cognitive flexibility was not found to be related to progression in analogy problem solving or training benefits. In addition, metacognition was revealed to be associated with training benefits. Implications for educational practice were provided in the discussion.

KEYWORDS
dynamic testing, executive functioning, giftedness

1 | INTRODUCTION

It has been proposed that cognitive abilities play an important role in children's school performance. Both intelligence (Roth et al., 2015), and executive functions (e.g., Monette, Bigras, & Guay, 2011; Viterbori, Usai, Traverso, & De Franchis, 2015) have been shown to predict school success. When a child is considered to be gifted in an educational context, this is often based on the results of an assessment procedure, including conventional, static testing of intelligence, or school aptitude. These tests, however, have been shown not to be advantageous for all children, and do not unveil information

about psychological processes involved in learning (e.g., Grigorenko, 2009). As conventional tests, for a large part, rely on past learning experiences (Elliott, Grigorenko, & Resing, 2010), children who have had less than favorable learning experiences, have been documented to underperform on these tests (Robinson-Zañartu & Carlson, 2013). Dynamic tests, in contrast, are much more focused on a child's potential for learning (Sternberg & Grigorenko, 2002). As in these tests, feedback and/or instruction are integrated into the testing procedure (Elliott, 2003), they allow for examining to what extent children show improvement in performance after an intervention, and whether other cognitive factors, such as executive functions, play a role in learning. In the current study, dynamic testing principles were applied to investigate to what extent two aspects of executive functioning, cognitive flexibility and metacognition, would be related to static or dynamic progression in analogy problem solving of gifted and average-ability children.

1.1 | Dynamic testing

Rather than measuring the knowledge or skills a child has already mastered, dynamic testing focuses on what a child would achieve in a short time frame, and this assessment procedure is therefore expected to provide a more complete picture of a child's potential for learning (Elliott, 2003). The pretest training posttest design (Sternberg & Grigorenko, 2002) is a frequently used application of dynamic testing that allows for structured measuring of a child's learning progression. The graduated prompts technique (e.g., Campione & Brown, 1987) has been used successfully as a training intervention in combination with said design. In this training approach, children are provided with structured prompts each time they make a mistake in problem solving. In the current study, prompts were tailored to each individual problem to be solved, and became more specific gradually, ranging from metacognitive to cognitive prompts and modeling (Resing & Elliott, 2011).

Similar to static test scores, dynamic testing outcomes have shown that there are many individual differences between children; both in terms of the instruction they require to show learning progression, as well as in terms of the level of progression they show after training (e.g., Resing, 2013). Dynamic testing of children who have strong cognitive capacities, nevertheless, seems an area researched less intensively. In earlier studies, dynamic tests for this group of learners have predominantly been used as a means to identify giftedness in disadvantaged populations (e.g., Kirschenbaum, 1998), such as those who are economically disadvantaged (e.g., Borland & Wright, 1994). Previous research further indicates that gifted children not only have a cognitive advantage, but, more specifically, learn new skills faster, and are better at generalizing newly acquired knowledge (Calero, García-Martín, & Robles, 2011). The potential role of executive functioning in dynamic testing of this group of children has, however, not yet been examined abundantly.

Dynamic tests frequently utilize inductive reasoning tasks (e.g., Resing, 2013; Stevenson, Heiser, & Resing, 2013). Inductive reasoning is believed to play a central role in intelligence (Klauer & Phye, 2008), and is said to be of crucial importance with regard to acquiring and applying knowledge (Goswami, 2012) and solving problems (Richland & Burchinal, 2012).

1.2 | Executive functioning

The graduated prompts technique employed in the current study included prompts activating different aspects of executive functioning, for example, in relation to self-regulation and monitoring of the problem-solving process. Executive functions comprise a number of complex cognitive processes enabling conscious control of thought and action (Monette et al., 2011) that are critical to purposeful, goal-directed behavior (Arffa, 2007). They are seen as the cognitive component of self-regulation (Calkins & Marcovitch, 2010). Research suggests that executive functions include inhibition, working memory and cognitive flexibility, which are key components of higher-order executive functions, such as metacognition (Miyake et al., 2000). The latter is usually divided into two dimensions: knowledge and regulation of cognitive activity (Schneider, 2010). To apply metacognition, assumed to play a role in developing new expertise (e.g., Sternberg, 1998), cognitive flexibility, working memory, and sufficient inhibition are prerequisites (Roebers, Cimeli, Röthlisberger, & Neuenschwander, 2012).

In addition, it has been argued that flexibility in applying newly learned skills and knowledge can be seen as an important aspect of cognitive functioning (e.g., Resing, 2013). Cognitive flexibility is said to include the ability to change perspectives spatially, or interpersonally, and being sufficiently flexible to adjust thinking to changing demands. Further, it is seen as a key component of the ability to think outside the box, and shares many characteristics with creativity, task, and set switching (Diamond, 2013).

Executive functioning has been found to be related to cognition (e.g., Ardila, Pineda & Rosselli, 2000). Studies investigating the relationship of executive functioning in a dynamic testing context, in particular with gifted children, however, are few, with most studies focusing on the role of working memory (e.g., Resing, Xenidou-Dervou, Steijn, & Elliott, 2012; Swanson, 2011).

1.3 | The current study

The current study utilized a dynamic test for analogical problem solving, a subtype of inductive reasoning, employing graduated prompts techniques. Our main research aim was to provide more insight into the potential benefits of dynamic testing of gifted children. More specifically, we focused on the roles that ability, cognitive flexibility, and metacognition play in repeatedly measured static versus dynamic progression in solving analogies.

Our first cluster of research questions addressed children's progression in solving analogies from pretest to posttest. Based on previous research into progression of unprompted solving of analogy problems among young children (e.g., Tunteler, Pronk, & Resing, 2008), we expected a significant main effect of time. We hypothesized (1a) that both unguided practice and dynamic testing would lead to progression in solving analogies from session to session. More importantly, we expected a significant interaction of time × condition, hypothesizing (1b) that children in the dynamic testing condition would show more progression from pretest, before training, to posttest, after training (e.g., Resing & Elliott, 2011; Stevenson et al., 2013). We further expected a significant interaction between time and ability. Gifted children were reported to have a more extensive zone of proximal development (e.g., Calero et al., 2011), therefore we hypothesized (1c) that gifted children would show more progression after unguided practice experiences than their average-ability peers. We also expected a significant interaction of time × condition × ability, indicating that gifted children would show more progression after training than their average-ability peers (1d).

Our second cluster of research questions concerned the association between executive functioning and children's progression from pretest to posttest. We expected a significant interaction between time and cognitive flexibility. Considering that flexibility in applying skills and knowledge is suggested to be important for learning and applying new knowledge (e.g., Resing, 2013), we hypothesized (2a) that children with higher levels of cognitive flexibility would show more progression in solving analogies than their peers with lower levels of cognitive flexibility. We also expected an interaction between time, condition, and cognitive flexibility, (2b) hypothesizing that children with higher levels of cognitive flexibility would benefit more from dynamic training than those with lower levels. Furthermore, a significant interaction between time, condition, ability, and cognitive flexibility was expected. Building on empirical studies in which high-ability children were found to have an advantage in executive functioning (e.g., Arffa, 2007), we hypothesized (2c) that the progression paths of gifted children with higher levels of cognitive flexibility would be steeper than those of their average-ability peers with similar levels of cognitive flexibility.

Moreover, as self-regulating, metacognitive skills were found to play a significant role in learning (e.g., Campione, Brown, & Ferrara, 1982; Sternberg, 1998), we expected an interaction between time and metacognition, hypothesizing (3a) that children with higher levels of metacognition would show more progression in solving analogies than their peers with lower levels of metacognition. We also expected a significant interaction between time, metacognition, and condition, and hypothesized (3b) that children with higher levels of metacognition would benefit more from training than their age mates with lower levels of metacognition. Finally, a significant interaction was expected between time, condition, ability, and metacognition. Taking into account that high-ability children were found to have an advantage in

TABLE 1 Overview of the hypotheses (SA = solving analogies)

	Hypothesis
1a	Unguided practice and dynamic testing will lead to progression in SA over time
1b	Dynamic testing will lead to more progression from pre- to posttest
1c	Gifted children will show more progression after unguided practice
1d	Gifted children will show more progression after training
2a	Higher levels of cognitive flexibility will lead to more progression in SA
2b	Higher levels of cognitive flexibility will lead to more progression after dynamic training
2c	Progression paths of gifted children with higher levels of cognitive flexibility will be steeper
3a	Higher levels of metacognition will lead to more progression in SA
3b	Children with higher levels of metacognition will benefit more from training
3c	Progression paths after training of the gifted children with higher levels of metacognition will be steeper
4a	Gifted children will need less metacognitive prompts
4b	Gifted children will need less cognitive prompts

self-regulation (e.g., Calero, García-Martín, Jiménez, Kazén, & Araque, 2007), we hypothesized (3c) that the progression paths after training of the gifted children who have higher levels of metacognition would be steeper than their average-ability peers with similar levels of metacognition.

Our last research question focused on individual differences in instructional needs, as measured by the number and the type of prompts required during training. As high-ability children were found to be more responsive to feedback (Kanevsky & Geake, 2004), and to have an advantage in self-regulation (e.g., Calero et al., 2007), we expected that gifted children's instructional needs during dynamic training would be significantly different from their average-ability peers. We hypothesized that gifted children would (4a) need both less metacognitive and (4b) less cognitive prompts than their average-ability peers. Table 1 provides an overview of the hypotheses.

2 | METHOD

2.1 | Participants

In this study, 113 children, 54 boys and 59 girls, participated, ranging in age from 7.1 to 8.9 years ($M = 7.90$). The average-ability children ($n = 68$) attended mainstream elementary schools, and those who were identified as gifted were enrolled in special settings for gifted and talented children in the Netherlands. Gifted children ($n = 45$) were oversampled and preliminary identification of giftedness took place on the basis of their enrolment in gifted education and qualitative judgments of parents and teachers regarding their giftedness.[1] Schools participated on a voluntary basis, and written permission to participate was obtained from the children's parents and schools prior to participation. Six children dropped out, as they did not participate in each test session.

2.2 | Design

The study utilized a 2×2 pretest–posttest control group design with randomized blocks with Ability category (gifted vs. average ability) and Condition (dynamic testing vs. unguided practice) as variables (see Table 2). Blocking was based on the scores on the Raven Standard Progressive Matrices test (Raven, 1981), administered before the pretest. All the children who had been identified as gifted had obtained Raven scores of at least the 90th percentile. Children in the dynamic testing subgroups received training between pretest 2 and posttest, whereas children in the unguided practice subgroups received an unrelated dot-to-dot control task of equal length between pretest 2 and posttest.

TABLE 2   Overview of the design

		Dynamic Testing (n = 53)		Unguided Practice (n = 60)	
		Gifted (n = 22)	Average Ability (n = 31)	Gifted (n = 23)	Average Ability (n = 37)
Prior to dynamic/static testing	Raven, BRIEF, BCST-64	x	x	x	x
Dynamic/static test	Pretests 1 and 2	x	x	x	x
	Dynamic training	Dynamic training	Dynamic training	Dot-to-dots control task	Dots-to-dots control task
	Posttest	x	x	x	x

The design included pretest sessions 1 and 2 to enable comparisons between static and dynamic progression. During the pretest sessions and the posttest, all children were only provided with short, general instructions. Administration of the instruments, including the training session, took approximately 20–30 minutes per session.

2.3 | Materials

2.3.1 | Raven

Participants were administered the Raven Standard Progressive Matrices Test (Raven, 1981) as a measure of intellectual ability and a blocking instrument. The Raven test is a nonverbal intelligence test that measures fluid intelligence by means of multiple choice figural analogies. The Raven Standard Progressive Matrices (internal consistency $r = .91$) consists of five sets of twelve items each, with a total of 60 items. In this study, only the raw scores were used in the analyses.

2.3.2 | Berg Card Sorting Test-64 (BCST-64)

The BCST-64 (Piper et al., 2011), the shortened version of the BCST, containing 64 trials, was used to measure cognitive flexibility. The BCST is an open-source computerized version of the Wisconsin Card Sorting Test (WCST; Grant & Berg, 1948). The unstandardized number of perseverative errors made during the administration of the BCST-64 were used as a measure of the participants' cognitive flexibility. Higher perseverative errors correspond with lower cognitive flexibility.

2.3.3 | BRIEF

The teacher questionnaire of the Dutch version of the Behavior Rating Inventory of Executive Functions (BRIEF; Smidts & Huizinga, 2009) was utilized to obtain teachers' evaluation of children's metacognition. The teacher questionnaire contains 86 items that make up eight scales, and two indices. Scores on the BRIEF Metacognition Index (Cronbach's $\alpha = .95$) were used to obtain the teacher's evaluation of each child's metacognition. Higher scores of the BRIEF are associated with more deviations from the norm, or impairment of executive functions. In the present study, raw scores were used.

2.4 | Dynamic version of geometric analogies

2.4.1 | Pretests and posttest

The dynamic test used in this study was composed of geometric visuospatial analogies of the type A:B::C:D (see Figure 1 for an example item). Both the pretests and the posttest consisted of 20 items of various difficulty. Six basic geometrical shapes were used in the construction of the analogies: squares, triangles, hexagons, pentagons, circles, and
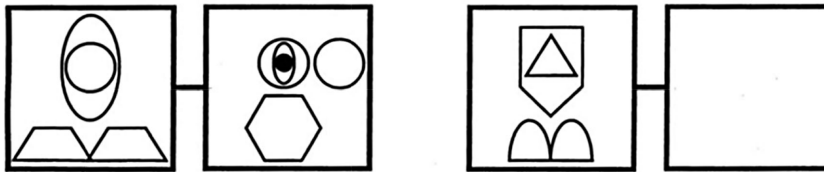
FIGURE 1 Example of a difficult analogy item

ovals. Each analogy was constructed by means of five possible transformations: changing position, adding or subtracting an element, changing size, halving, and doubling. The test was administered as an open-ended paper-and-pencil test, and children had to draw their answers.

The pretests and posttest, parallel sessions with different, but equivalent analogy items, were composed of 20 trials. The test sessions were equivalent in terms of the numbers of different elements, and transformations used for each analogy item, as well as the order in which the items were presented in relation to their difficulty level. The children received minimal instructions only in the two pretests and the posttest, as they were told that they had to solve puzzles with different shapes. The test leader then asked the child which shapes had to be drawn in the fourth box to solve the puzzle.

2.4.2 | Training

The current study employed one training session, consisting of 10 geometric analogies that were not used in either the pretests or the posttest. The training session was based on graduated prompts techniques (Campione & Brown, 1987; Resing & Elliott, 2011), and consisted of five steps per item. The prompts were administered following a standardized protocol, and were provided hierarchically, from two very general metacognitive prompts to two concrete cognitive prompts tailored to each specific item (see Appendix Table A1) . Prompts were given if a child could not solve the analogy independently. After each prompt, children were asked to draw the solution of the analogy, and check their answer. If, after the fourth prompt, a child had not solved the analogy correctly, the test leader modeled the correct answer for the child. After the four prompts had been provided, and/or the test leader had shown the correct answer, the children were asked to explain why they thought their answer was correct. Then, the test leader provided a correct self-explanation.

2.5 | General procedure

The children were tested once a week over a period of five consecutive weeks. All tests and questionnaires part of the present study were administered following standard, protocolled instruction. At the beginning of the pretests, training session, and posttest, the children were provided with the six geometrical shapes used in the analogies, and in cooperation with the test leader named each shape, after which the test leader asked the child to draw the shapes below the printed shapes, staying as close to the original as possible.

2.6 | Scoring

Analogy items were scored on the basis of children's drawings, in combination with their verbal explanations. Some of the children experienced difficulties drawing the geometrical shapes. As each child had to copy the shapes used in the analogies on the cover sheet, in the vast majority of cases the test leader knew which shapes the child was drawing. If necessary, the child would be asked to point out on the cover sheet which shapes were intended.

For each item, the number of transformations that the child had applied correctly in solving the analogy was scored. Each analogy item was constructed by means of 1, 2, 3, 4, or 6 transformations that the child had to apply correctly to accurately solve the item, adding up to a total of 59 transformations per test session. The total number of transformations applied correctly in solving the analogies was taken as the outcome variable for each test session.

To estimate coding reliability, the pretest 1 data were scored by both the first author and a student assisting in data collection. An inter-rater reliability analysis showed that inter-rater agreement for the pretest 1 correct transformations was good ($\kappa = .83$, $p < .0001$).

2.7 | Analyses

Multilevel modeling was used to analyze the data. Multilevel modeling capitalizes on the hierarchical structure of the data, allowing us to study relations among variables at different levels and across levels. We can simultaneously answer level 1 questions about within-person change, and level-2 questions about how these changes vary across children (Singer & Willett, 2003). In the current study, level 1 represented the repeated measurements of the number of correct transformations within children, and level 2 represented the variability between children. We followed a predetermined model building structure as proposed by Singer and Willett (2003); starting with two simple, unconditional models and including our time-variant and time-invariant predictors in the successive models. The predictors were: condition, ability category, cognitive flexibility, and metacognition. Two time-invariant predictors, metacognition and cognitive flexibility, were mean centered to improve interpretation (Singer & Willett, 2003).

R (R Development Core Team, 2014) was used to fit the models. The fit of all models was compared using the likelihood ratio test (LRT) and two fit indices: Akaike's Information Criterion (AIC) and the Schwarz's Bayesian Information Criterion (BIC). The LRT follows a χ^2-distribution where the degrees of freedom are equal to the difference in the number of estimated parameters between the models. The LRT compares the "log likelihood" of two models and tests whether they differ significantly. The AIC and BIC are ad hoc criteria based on the log-likelihood statistic. The AIC and BIC statistics can be compared for all pairs of models, whether the models are nested within one another or not. These indices use a penalty function based on the number of parameters so that the more parsimonious model is favored. A lower AIC and BIC value indicates a better fit of the model (Singer & Willett, 2003). All the discussed models were fitted using the full maximum-likelihood (FML) estimation. Most of the models differed in their fixed parts, and therefore deviance based on FML was needed to be able to compare the successive models (Singer & Willett, 2003).

3 | RESULTS

Before examining our research questions, one-way analyses of variance were conducted for each Ability category to evaluate possible differences between children in the conditions. The Raven scores, pretest 1 number of correct transformations, and age in months were used as dependent variables, and Condition (dynamic testing vs. unguided practice) as independent variable. The findings for the gifted and average-ability children, analyzed separately, revealed no significant differences in Raven scores ($p = .53$; $p = .61$), pretest 1 correct transformations ($p = .40$; $p = .85$), nor in age ($p = .52$; $p = .98$) between the dynamic testing and unguided practice conditions, respectively. We also examined possible differences between the gifted and average-ability children. The gifted children outperformed their peers on both the Raven scores, and the pretest 1 correct transformations (for both measures, $p < .001$), but no significant differences were found in age ($p = .31$). Descriptive statistics of all measures used in the current study, per condition and Ability category are provided in Table 3.

We conducted growth curve analyses (Multilevel analysis; MLA) to model growth in the number of correct transformations. Table 4 presents the parameters and fit indices of the models. We first fitted the unconditional means model (intercept-only model) to acquire the random effects that revealed a significant intercept effect ($p < .001$). We examined the intraclass correlation coefficient (ICC) as a measure of dependence; it describes the proportion of outcome variance that lies between persons in the population (i.e., the cluster structure of the data). As indicated by the ICC coefficient, of the total variation in the number of correct transformations, 54.38% could be attributable to differences between children. This finding revealed that the observations were not independent, and indicated that there was systematic variation in the outcome measure (transformations) worth exploring, both for the within-level and between-level variance, reinforcing the choice of multilevel modeling.

TABLE 3 Mean scores and standard deviations of Raven scores, pretest 1, pretest 2, posttest correct transformations, cognitive flexibility, and metacognition per condition and ability group

		Gifted		Average Ability	
		Dynamic Testing	**Unguided Practice**	**Dynamic Testing**	**Unguided Practice**
	N	22	23	31	37
Raven (raw scores)	M (SD)	43.82 (4.22)	44.57 (3.78)	34.55 (5.53)	33.78 (6.47)
Pretest 1 (transformations)	M (SD)	39.14 (15.13)	41.96 (9.26)	29.16 (13.56)	28.43 (15.77)
Pretest 2 (transformations)	M (SD)	46.86 (17.62)	53.74 (4.05)	43.52 (13.40)	41.03 (18.27)
Posttest (transformations)	M (SD)	54.59 (9.63)	53.91 (5.97)	52.77 (7.14)	41.68 (18.14)
Cognitive flexibility (perseverative errors)	M (SD)	11.36 (5.14)	12.87 (7.43)	9.81 (5.53)	13.84 (7.79)
Metacognition (raw scores)	M (SD)	59.91 (15.68)	61.61 (20.28)	59.47 (17.21)	60.30 (15.42)

TABLE 4 Results of the fitted multilevel models for the number of correct transformations

Model	Estimate(SE)	Deviance	AIC	BIC
1. Intercept only	42.89(1.26)**	2,750.6	2,756.6	2,768.1
2. Time	8.13(.51)**	2,557.8	2,569.8	2,592.7
3. Condition	3.51(1.40)*	2,552.3	2,566.3	2,593.1
4. Ability category	8.23(2.39)**	2,541.5	2,557.5	2,588.1
5. Ability category × Time	−2.21(.98)*	2,536.5	2,554.5	2,589.0
6. Ability category × Condition	−3.85(2.82)	2,534.8	2,554.8	2,593.1
7. Cognitive flexibility	−.13(.17)	2,536.0	2,556.0	2,594.3
8. Cognitive flexibility × Time	.02(.07)	2,536.0	2,558.0	2,600.0
9. Cognitive flexibility × Condition	.34(.21)	2,533.7	2,555.7	2,597.8
10. Cognitive flexibility × Condition × Ability category	.49(.35)	2,534.1	2,556.1	2,598.2
11. Metacognition	−.03(.07)	2,513.7	2,533.7	2,571.9
12. Metacognition × Time	.05(.03)	2,510.8	2,532.8	2,574.8
13. Metacognition × Condition	.15(.07)*	2,509.3	2,531.3	2,573.3
14. Metacognition × Condition × Ability category	−.06(.14)	2,509.1	2,533.1	2,578.9

Note. Significance: **$p < .001$, *$p < .05$. The deviance, AIC, and BIC statistics were examined for the relative goodness of fit of the successive models.

In Model 2 (the unconditional growth model), we included our time predictor into the level-1 submodel to explain the remaining within-child variance (117.8). The estimated rate of change in the number of correct transformations for an average participant was 8.13 ($p < .001$); children generally improved in the number of correctly applied transformations. A negative covariance (−.56) was found between the slope and intercept. This indicated that children using fewer correct transformations at pretest 1 increased their number of correct transformations slightly faster across test sessions than children with a higher number of correct transformations at pretest 1. Variance components revealed remaining variance in the number of correct transformations both between, and within, children. Extending the model by adding other predictors could possibly reduce this variation.

Model 3 included Condition as an explanatory variable for the number of correct transformations. Result of the LRT showed that model fit improved ($\chi^2(1) = 5.46$, $p = .02$). Children of the unguided practice group had, on average, an estimated rate of change of 7.31. Therefore, these children generally increased their number of correct transformations across test sessions. A positive fixed effect for Condition (training vs. unguided practice) of 3.51 revealed that the dynamic training session influenced the performance of the children. In accordance with our expectation, those who

received a dynamic training session improved more in the number of correct transformations from pretest 2 to posttest than the children in the unguided practice condition.

In Model 4 we included Ability category, gifted versus average-ability, as a predictor for initial status. Model 4 provided a better fit to the data compared to Model 3 ($\chi^2(1) = 10.82, p = .001$). Children's Ability category was found to be related to the number of correct transformations at pretest 1 as shown by a significant main effect of Ability category (8.23). Specifically, children with higher intellectual ability scored, on average, higher on pretest 1 than average-ability peers. Model 5 showed that Ability category was also a significant predictor for children's rate of change, as indicated by a significant interaction of Ability category and Time. Model fit improved ($\chi^2(1) = 4.96, p = .03$). The estimate ($-2.21$) revealed that average-ability children improved more in the number of correct transformations over time than gifted children.

In Model 6 we examined whether the dynamic training session had different benefits for gifted and average-ability children. We included the interaction effect of Ability category and Condition, which did not improve model fit ($\chi^2(1) = 1.75, p = .19$). No significant difference was found in dynamic training benefits for gifted and average-ability children, as revealed by the nonsignificant interaction effect (-3.85), indicating that gifted children did not show more progression in the number of correct transformations after training than their average-ability peers.

Model 7 showed no significant main effect of Cognitive flexibility; model fit did not improve ($\chi^2(1) = .53, p = .47$). The nonsignificant interaction effect of Cognitive flexibility × Time in Model 8 ($\chi^2(2) = .59, p = .75$) indicated that we could not support our expectation that children with higher levels of cognitive flexibility would show more progression in the number of correct transformations than their age mates with lower levels of cognitive flexibility. Children with higher levels of cognitive flexibility did also not benefit more from the dynamic training session than children with lower levels of cognitive flexibility as shown in Model 9 ($\chi^2(2) = 2.84, p = .24$). Furthermore, results of Model 10 showed that the progression paths of gifted children that had higher levels of cognitive flexibility were not steeper than those of their average-ability peers ($\chi^2(2) = 2.47, p = .29$). The time-invariant predictor Cognitive flexibility was not included in the remaining models.

Model 11 included the main effect of Metacognition. A nonsignificant effect was found, however, model fit did improve after inclusion of the predictor ($\chi^2(1) = 22.80, p < .001$). Results of Model 12 showed that children with higher scores on the Metacognition Index showed equivalent progression in the number of correct transformations across test sessions than their peers with lower scores on the Metacognition Index ($\chi^2(1) = 2.97, p = .08$). In Model 13, we included the interaction effect of Metacognition and Condition, which led to an improvement in model fit ($\chi^2(1) = 4.40, p = .04$). The estimate (.149) showed that children with higher scores on the Metacognition Index benefited more from training than peers with lower scores. We included the three-way interaction between Condition, Ability category, and Metacognition in Model 14. Results showed that the progression paths of gifted children that had higher levels of metacognition were not steeper than those of their average-ability peers ($\chi^2(1) = .20\ p = .66$).

In conclusion, Model 13 was shown to be the model that best fitted the data based on the LRT, and the AIC and BIC statistics. The dynamic sessions led to an improvement in the number of correct transformations the children used. No differences in dynamic training benefits for gifted and average-ability children were found. The average-ability children in the unguided practice condition did, however, show more improvement across test sessions than the gifted children in the unguided practice session. Cognitive flexibility did not influence children's progression over time and the improvement in the number of transformations after receiving the dynamic training. The progression paths did also not differ for gifted children with higher levels of cognitive flexibility and their average-ability peers. Metacognition did not influence progression in the number of correct transformations. Children with lower levels of metacognition, as indicated by higher scores on the Metacognition Index, showed more improvement in the number of correct transformations after the dynamic training than their peers with higher levels of metacognition. Lastly, the progression paths did not differ between gifted children who had higher levels of metacognition and their average-ability peers.

To examine our final research question regarding potential differences in the instructional needs of gifted and average-ability children, we conducted a one-way analysis of variance (ANOVA) with two within-subjects factors (metacognitive and cognitive prompts) and one between-subjects (Ability category) factor with the number of prompts

TABLE 5 Mean scores and standard deviations of the number of metacognitive and cognitive prompts received during training per Ability category

	Metacognitive Prompts			Cognitive Prompts	
	M	SD		M	SD
Gifted	11.91	2.14		2.41	4.47
Average ability	12.87	2.39		2.90	4.29

in each category as dependent variables. No significant differences were found in the number of metacognitive, $F(1,51)$ = 2.27, p = .14, or cognitive prompts, $F(1,51)$ = .17, p = .69 across ability categories (see Table 5). These results suggested that the two groups of children, gifted versus average-ability needed a similar number of steps during training, indicating their need for instruction was similar from both a quantitative, relating to the total number of prompts, and a qualitative, relating to differences in the type of prompts provided, perspective.

4 | DISCUSSION

The current study explored the potential differential benefits of dynamic versus static testing of gifted and average-ability children, and focused on two aspects of executive functioning, cognitive flexibility and metacognition. First of all, our results showed that children who had unguided practice experience only, and children who were dynamically tested showed progression in the number of correct analogical transformations. When children were tested dynamically, however, their progression paths were shown to be more advanced, which supports previous findings (e.g., Stevenson et al., 2013). In this sense, our findings build upon earlier studies in which it was posited that dynamic testing of children reveals a more complete picture of their cognitive potential than static testing only (e.g., Elliott, 2003).

Moreover, our findings indicated, as expected, that gifted children start at a higher ability point, and keep this advantage during following sessions. When looking into potential differences between gifted and average-ability children in relation to the nature of progression, in contrast to our expectations, it was found that, in general, the average-ability children showed more progression than their gifted peers. We cannot, however, discount that the gifted children in the current study might have experienced a ceiling effect in testing. If so, we would then have expected them to show a differential need for instructions, which could not be supported by our data. Moreover, no mention of a ceiling effect is made in previous research with participants of the same age (e.g., Tunteler et al., 2008). It must be mentioned, nevertheless, that it is not known whether any high-ability children participated in these studies. Therefore, this explanation requires further research.

Looking more closely into training benefits, it was revealed that the gifted and average-achieving children showed similar rather than different progression lines after training, whereas previous studies into dynamic testing of gifted children found that these groups of children differed significantly in their performance and progression (e.g., Calero et al., 2011; Kanevsky & Geake, 2004). In the light of the fact that all groups of children progressed after training, our findings, ultimately, seem to suggest that dynamic testing might be better suited to reveal the cognitive potential of all groups of children (Elliott et al., 2010), including those with above-average cognitive abilities.

We also examined the role that cognitive flexibility and metacognition play in progression in accuracy of analogical reasoning, and training benefits. It could not be established that cognitive flexibility plays a role. A number of reasons can be identified for the unexpected results regarding cognitive flexibility. First of all, research into executive functioning among children is challenging. One important reason is the type of instruments used to measure executive functioning. It has been noted that performance-based tasks, such as the BCST-64 used in the current study, rarely measure one executive function only (e.g., Miyake et al., 2000). By definition, executive functions regulate various cognitive processes, including for instance visuospatial processing. Performance-based tasks measure these other processes as well, making measuring just one executive function, in isolation, difficult (Viterbori et al., 2015). The

developmental nature of executive functions in childhood should also be taken into consideration (e.g., Diamond, 2013). Moreover, it should be noted that the cognitive flexibility task used in the current study is a single measurement, static test, whereas learning potential measures are dynamic. Therefore, future studies could research this relationship further by utilizing a dynamic cognitive flexibility task, such as the dynamic Wisconsin Card Sorting Task (e.g., Boosman, Visser-Meily, Ownsworth, Winkens, & Van Heugten, 2014). These authors found that the dynamic executive functioning indices were significantly associated with cognitive functions, whereas the static indices were not.

It was, nonetheless, found that metacognition had an effect on the training benefits, but not on the progression from pretest to posttest. Children who, according to their teachers, had lower levels of metacognition, in contrast with our expectations, benefitted more from training than their peers with higher levels of metacognition. Furthermore, the findings provide a first indication that a graduated prompts training procedure can, to a certain extent, compensate for lower levels of metacognition. This notion is particularly relevant considering Sternberg's (1998) assertion that metacognition is an important ability in the development of expertise.

Although it seems plausible that the graduated prompts technique used in the current study also helps improve metacognition, this tentative hypothesis should be investigated using several measurements of metacognition. It must be noted that, although studies suggest that rating scales can be used successfully to obtain an approximation of children's executive functioning (Toplak, West, & Stanovich, 2013), using teacher ratings is a very indirect method of measuring metacognition. However, due to the young age of the participants, it was not possible to use other instruments to obtain metacognition measures. Self-report measures are not recommended for young children, as they rely heavily on verbal ability (Whitebread et al., 2009). Thinking aloud protocols, moreover, might not fully capture implicit cognitive processes, as young participants might not be conscious of their metacognitive processes while solving a task (Lai, 2011). In future research among older children, these instruments could be used to investigate the relationship between metacognition and dynamic testing measures. Future studies should also focus on development and implementation of instruments that directly measure or predict executive functioning among young children.

Finally, we looked more closely into children's instructional needs during dynamic training. Contrary to what we expected based on previous literature (e.g., Calero et al., 2007; Kanevsky & Geake, 2004), we found no differences in the instructional needs of the gifted versus average-ability groups of children: the two groups of children needed a similar number of cognitive and metacognitive prompts. These results ultimately suggest that, compared with their average-ability peers, gifted children did not differ in terms of the number of cognitive, metacognitive prompts, nor in the extent to which they needed modeling, and, thus, can have similar needs for instructions to progress in learning. Individual differences between children's need for instructions, both within and across ability categories, were, however, found, as suggested by the standard deviations of both groups of children, which is in line with previous studies (e.g., Resing, 2013).

In addition to the limitations mentioned above, the current study encountered some other limitations. First of all, it is important to mention that we only used the Raven Standard Progressive Matrices as a measure of intellectual ability. Although the Raven test is known as a robust measure of intellectual ability (e.g., Jensen, 1998), we did not include other factors deemed important for cognitive and intellectual functioning, such as task commitment or creativity (e.g., Renzulli & D'Souza, 2014). Moreover, we only investigated correct analogical transformations, while other factors have also been shown to be important in progression in analogical reasoning. Investigating strategy use, in particular, could lead to interesting findings considering the assumed relationship between strategy use and aspects of executive and intellectual functioning (e.g., Shore, 2000).

The results of the current study yield some important implications for educational professionals. It seems advisable to administer a dynamic rather than a static test when children's intellectual abilities are questioned, especially for children with lower levels of metacognition. In this light, investigating the interrelationship between executive functioning and dynamic testing seems worthwhile, especially for children with lower levels of intellectual functioning or learning disabilities. The benefits of dynamic testing for these special groups of children seem especially relevant within the framework of response to intervention (RTI; e.g., Grigorenko, 2009). Research suggests dynamic testing may be used successfully to identify or predict the responsiveness to intervention of these children (e.g., Fuchs, Compton, Fuchs, Bouton, & Caffrey, 2011).

Opponents of dynamic testing often argue that testing dynamically is more labor intensive, and, thus, more expensive than testing statically. The dynamic test used in the present study, for example, in total, took approximately 60–90 minutes to administer, whereas for a static test with a single test session, 15–20 minutes would suffice. Nevertheless, our findings suggest that taking extra time to test these children, including those identified as gifted, more than once and administering a dynamic training session, helps them in unveiling their cognitive abilities, and, thus, is worth the extra investment.

This notion becomes even more salient when taking into account that dynamic testing of children also provides insight into their instructional needs (e.g., Bosma & Resing, 2012). The results of the current study remind us that, when teaching high-ability children, these children do not, by definition, need less instruction or feedback than average-ability children, to show progression in learning. Just like any other children, some of these children can also profit from extra feedback or help so they can unveil their true cognitive potential. Finally, and most importantly, the results of the present study indicate that children, even those who have already achieved excellent results, can show learning progression when they are provided with the right instructions.

ENDNOTE

[1] In the Netherlands, intelligence testing is not standard practice in primary schools. For admittance to special talent or gifted educational programs, teachers and parents' nominations are often used. In the present study, these nominations were used, in combination with a percentile rank score of at least 90 to identify children as gifted.

REFERENCES

Ardila, A., Pineda, D., & Rosselli, M. (2000). Correlation between intelligence test scores and executive function measures. *Archives of Clinical Neuropsychology*, *150*, 31–36. https://doi.org/10.1093/arclin/15.1.31

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

Boosman, H., Visser-Meily, J. M. A., Ownsworth, T., Winkens, I., & Van Heugten, C. M. (2014). Validity of the Dynamic Wisconsin Card Sorting Test for assessing learning potential in brain injury rehabilitation. *Journal of the International Neuropsychological Society*, *20*, 1034–1044. https://doi.org/10.1017/s1355617714000897

Borland, J. H., & Wright, L. (1994). Identifying young, potentially gifted, economically disadvantaged students. *Gifted Child Quarterly*, *38*, 164–171. https://doi.org/10.1177/001698629403800402

Bosma, T., & Resing, W. C. M. (2012). Need for instruction: Dynamic testing in special education. *European Journal of Special Needs Education*, *27*, 1–19. https://doi.org/10.1080/08856257.2011.613599

Calero, M. D., García-Martín, M. B, Jiménez, M. I., Kazén, M., & Araque, A. (2007). Self-regulation advantage for high-IQ children: Findings from a research study. *Learning and Individual Differences*, *17*, 328–343. https://doi.org/10.1016/j.lindif.2007.03.012

Calero, M. D., García-Martín, M. B., & Robles, M. A. (2011). Learning potential in high IQ children: The contribution of dynamic assessment to the identification of gifted children. *Learning and Individual Differences*, *21*, 176–181. https://doi.org/10.1016/j.lindif.2010.11.025

Calkins, S. D., & Marcovitch, S. (2010). Emotion regulation and executive functioning in early development: Integrated mechanisms of control supporting adaptive functioning. In S. D. Calkins & M. A. Bell (Eds.), *Child development: At the intersection of emotion and cognition* (pp. 37–58). Washington, DC: APA Press.

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: Guilford Press.

Campione, J. C., Brown, A. L., & Ferrara, R. A. (1982). Mental retardation and intelligence. In R. J. Sternberg (Ed.), *Handbook of human intelligence* (pp. 392–490). New York, NY: Cambridge University Press.

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

Elliott, J. G. (2003). Dynamic assessment in educational settings: Realising potential. *Educational Review*, *55*, 15–32. https://doi.org/10.1080/00131910303253

Elliott, J. G., Grigorenko, E. L., & Resing, W. C. M. (2010). Dynamic assessment: The need for a dynamic approach. In P. Peterson, E. Baker, & B. McGaw (Eds.), *International encyclopedia of education* (Vol. 3, pp. 220–225). Oxford: Elsevier. https://doi.org/10.1016/b978-0-08-044894-7.00311-0

Fuchs, D., Compton, D. L., Fuchs, L. S., Bouton, B., & Caffrey, E. (2011). The construct and predictive validity of a dynamic assessment of young children learning to read: Implications for RTI frameworks. *Journal of Learning Disabilities*, 44, 339–347. https://doi.org/10.1177/0022219411407864

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/Dordrecht/Heidelberg/London: Springer. https://doi.org/10.1007/978-1-4419-1428-6_993

Grant, D. A., & Berg, E. A. (1948). A behavioral analysis of degree of reinforcement and ease of shifting to new responses in a Weigl-type card sorting problem. *Journal of Experimental Psychology*, 34, 404–411. https://doi.org/10.1037/h0059831

Grigorenko, E. L. (2009). Dynamic assessment and response to intervention: Two sides of one coin. *Journal of Learning Disabilities*, 42(2), 111–132. https://doi.org/10.1177/0022219408326207

Jensen, A. R. (1998). *The g factor: The science of mental ability*. Westport, CT: Praeger.

Kanevsky, L. S., & Geake, J. (2004). Inside the zone of proximal development: Validating a multifactor model of learning potential with gifted students and their peers. *Journal for the Education of the Gifted*, 28, 182–217. https://doi.org/10.1177/016235320402800204

Kirschenbaum, R. J. (1998). Dynamic assessment and its use with underserved gifted and talented populations. *Gifted Child Quarterly*, 42, 140–147. https://doi.org/10.1177/001698629804200302

Klauer, K. J., & Phye, G. D. (2008). Inductive reasoning: A training approach. *Review of Educational Research*, 78, 85–123. https://doi.org/10.3102/0034654307313402

Lai, E. R. (2011). Metacognition: A literature review. Always learning: Pearson research report.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology*, 41, 49. https://doi.org/10.1006/cogp.1999.0734

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

Piper, B. J., Li, V., Eiwaz, M. A., Kobel, Y. V., Benice, T. S., Chu, A. M., … Mueller, S. T. (2011). Executive function on the psychology experiment building language tests. *Behavior Research Methods*, 44, 110–123. https://doi.org/10.3758/s13428-011-0096-6

R Development Core Team. (2014). *R: A Language and Environment for Statistical Computing (Version 3.12) [Software]*. Vienna, Austria: R Foundation for Statistical Computing. Retrieved on 2 January 2017 from http://www.R-project.org

Raven, J. (1981). *Manual for Raven's progressive matrices and vocabulary scales*. Oxford, UK: Oxford Psychologists Press.

Renzulli, J. S., & D'Souza, S. (2014). Intelligences outside the normal curve: Co-cognitive factors that contribute to the creation of social capital and leadership skills in young people. In J. A. Plucker & C. M. Callahan (Eds.), *Critical issues and practices in gifted education: What the research says* (pp. 343–362). Waco, TX: Prufrock Press.

Resing, W. C. M. (2013). Dynamic testing and individualized instruction: Helpful in cognitive education? *Journal of Cognitive Education and Psychology*, 12, 81–95. https://doi.org/10.1891/1945-8959.12.1.81

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., Xenidou-Dervou, I., Steijn, W. M. P., & Elliott, J. G. (2012). A "picture" of children's potential for learning: Looking into strategy changes and working memory by dynamic testing. *Learning and Individual Differences*, 22, 144–150. https://doi.org/10.1016/j.lindif.2011.11.002

Richland, L. E. & Burchinal, M. R. (2012). Early executive function predicts reasoning development. *Psychological Science, 24*, 87–89. https://doi.org/10.1177/0956797612450883

Robinson-Zañartu, C., & Carlson, J. (2013). Dynamic assessment. In K. F. Geisinger (Ed.), *APA handbook of testing and assessment in psychology* (Vol. 3, pp. 149–167). Washington, DC: American Psychological Association. https://doi.org/10.1037/14049-007

Roebers, C. M., Cimeli, P., Röthlisberger, M., & Neuenschwander, R. (2012). Executive functioning, metacognition, and self-perceived competence in elementary school children: An explorative study on their interrelations and their role for school achievement. *Metacognition and Learning, 2*, 151–173. https://doi.org/10.1007/s11409-012-9089-9

Roth, B., Becker, N., Romeyke, S., Schäfer, S., Domnick, F., & Spinath, F. M. (2015). Intelligence and school grades: A meta-analysis. *Intelligence*, 53, 118–137. https://doi.org/10.1016/j.intell.2015.09.002

Schneider, W. (2010). Metacognition and memory development in childhood and adolescence. In H. W. Waters, & W. Schneider (Eds.), *Metacognition, strategy use and instruction* (pp. 54–81). New York: Guildford Press.

Shore, B. M. (2000). Metacognition and flexibility: Qualitative differences in how gifted children think. In R. C. Friedman, & B. M. Shore (Eds.), *Talents unfolding: Cognition and development* (pp. 167–187). Washington, DC: American Psychological Association.

Singer, J. D., & Willett, J. B. (2003). *Applied longitudinal data analysis: Modeling change and event occurrence.* New York, NY: Oxford University Press.

Smidts, D. P., & Huizinga, M. (2009). *BRIEF Executieve Functies Gedragsvragenlijst: Handleiding [Manual for the BRIEF Executive Functions Behavioral Questionnaire].* Amsterdam: Hogrefe.

Sternberg, R. J. (1998). Metacognition, abilities, and developing expertise: What makes an expert student? *Instructional Science*, *26*, 127–140. https://doi.org/10.1023/a:1003096215103

Sternberg, R. J., & Grigorenko, E. L. (2002). *Dynamic testing. The nature and measurement of learning potential.* Cambridge: 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

Swanson, H. L. (2011). Dynamic testing, working memory and reading comprehension growth in children with reading disabilities. *Journal of Learning Disabilities*, *44*, 358–371. https://doi.org/10.1177/0022219411407866

Toplak, M. E., West, R. F., & Stanovich K. E. (2013). Practitioner review: Do performance-based measures and ratings of executive function assess the same construct? *Journal of Child Psychology and Psychiatry*, *54*, 131–143. https://doi.org/10.1111/jcpp.12001

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

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

Whitebread, D., Coltman, P., Pasternak, D. P., Sangster, C., Grau, V., Bingham, S., … Demetriou, D. (2009). The development of two observational tools for assessing metacognition and self-regulated learning in young children. *Metacognition and Learning*, *4*, 63–85. https://doi.org/10.1007/s11409-008-9033-1

APPENDIX

TABLE A1   Schematic overview of the graduated prompts training protocol

Step	Instruction	Incorrect Answer?	Correct Answer?
1	This is another puzzle with four boxes. Do you remember what we are going to do? [*have child provide an answer*] We are going to solve the puzzle by filling the empty box with the correct figures. Just draw the answer that you think is correct in the empty box [*have child draw the answer*]. Check whether you drew the correct answer [*have child check and correct answer if necessary*]	The picture you drew is great, but it is not entirely correct yet. I will help you, but try to find the correct answer with as little help from me as possible. We will start again after each try.	To step 5: Well done, that is the correct answer! Can you tell me why this this the correct answer? [*Test leader models correct self-explanation, as per the protocol, tailored to each item*]
2	How do we start? [*have child provide an answer*] First, have a good look at the figures in these three boxes [*point at A, B, C*] Do you now know the correct answer? Just draw the answer that you think is correct in the empty box [*have child draw the answer*] Check whether you drew the correct answer [*have child check and correct answer if necessary*]	Great picture! It is not entirely correct. I will help you some more.	
3	Have a good look at these boxes [*point at A and B]* What do you see? [*Have child provide an answer*] We see that A and B belong together. Do you know why? [*have child provide an answer*] [*Then explain the transformations from A → B according to protocol, tailored per item*] Do you now know the correct answer? Just draw the answer that you think is correct in the empty box [*have child draw the answer*] Check whether you drew the correct answer [*have child check and correct answer if necessary*]	You drew another beautiful picture. It is almost correct, so I will help you a little bit more.	
4	Now have a good look at this box [*point at C*] and this box [*point at A]* What do you see? [*Have child provide an answer*] We see that A and C look alike, but that they changed a little bit. Can you tell me why? [*Have child provide an answer*] [*Then explain the similarities between A and C, B according to protocol, tailored per item*] Do you now know the correct answer? Just draw the answer that you think is correct in the empty box [*have child draw the answer*] Check whether you drew the correct answer [*have child check and correct answer if necessary*]	What a beautiful picture. You can draw very well. It is not entirely correct; I will show you the correct answer [*test leader draws correct answer*] Can you tell me why this this the correct answer? [*Test leader models correct self-explanation, as per the protocol, tailored to each item*]	