Learning trajectories in analogical reasoning : exploring individual differences in children's strategy paths

Pronk, C.M.E.

Author: Pronk, C.M.E.
Title: Learning trajectories in analogical reasoning : exploring individual differences in children's strategy paths
Issue Date: 2014-02-19

CHAPTER 3

The influence of dynamic testing and working-memory capacity on children's analogical reasoning: A microgenetic investigation using multilevel analysis

Pronk, C.M.E., Elliott, J.G., & Resing, W.C.M. (Submitted). The influence of dynamic testing and working-memory capacity on children's analogical reasoning: A microgenetic investigation using multilevel analysis.
Acknowledgements: We are grateful for the helpful comments on task design and multilevel analysis that were provided respectively by Claire Stevenson, Mark de Rooij, and Rien der Leeden (deceased).

Abstract

In the current study we investigated the inter- and intra-individual developmental trajectories of analogical reasoning in a dynamic test and non-guided practice setting. The study employed a microgenetic research method together with Multilevel Analysis to investigate developmental trajectories as a function of their background variables and experimental treatment. Background variables included verbal and abstract-visual-spatial working-memory capacity. Participants were 32 children aged 7-8 years with a mean age of 90 months. Half of the children followed a microgenetic design; the others followed a comparable design but were dynamically tested halfway the experiment, all assessment moments involving solving visual-spatial analogies tasks. All test sessions were undertaken individually. After repeated assessment sessions, children showed inter-individual growth in analogical reasoning through non-guided practice, but even more through dynamic testing. Growth through both practice and dynamic testing appeared to be influenced by spatial working-memory capacity. After dividing children into subgroups, multilevel analysis allowed us to display intra-individual developmental trajectories that were similar in amount and rate of analogical reasoning change within each subgroup. These study outcomes suggest a need for more in-depth microgenetic research into dynamic testing of analogical reasoning in combination with working-memory assessment. In particular, comparing the strategy use of subgroups painted by the current study might be very promising in revealing specific strengths and weaknesses that influence particular learning trajectories. This information could be used to better predict and ameliorate children's projected learning trajectories.

3.1 Introduction

The development and training of inductive reasoning, particularly children's capacity to reason by analogy, have been the focus of much research (e.g., Alexander, Willson, White, & Fuqua, 1987; Alexander et al., 1989; Goswami, 1992; Resing, 2000; Tzuriel & George, 2009). In former studies, children older than 6 years have typically displayed clear improvements in analogical reasoning after receiving a brief period of training or, alternatively, after having been given extensive instructions for such tasks as verbal analogies (Resing 1993, 1997), physical problem analogies (Tunteler & Resing, 2007a), concrete pictorial analogies (Hessels-Schlatter, 2002; Schlatter & Büchel, 2000; Stevenson, Resing, & Froma, 2009), and classic geometric analogies (Hosenfeld, Van der Maas, & Van den Boom, 1997a, 1997b; Tunteler, Pronk, & Resing, 2008). In contrast, younger children have tended only to show such gains when they had received extensive training (e.g., Tunteler & Resing, 2007a; Alexander et al., 1989). Therefore, this study focused on grade two children to investigate the development of analogical reasoning as it happens. This form of reasoning was induced by repeated non-guided practice and the use of a dynamic test employing concrete, figural analogies.

Both repeated non-guided practice, and instruction while learning, have been recognized as valuable in investigating developmental trajectories by means of a microgenetic research design (Winne & Nesbit, 2010; Siegler, 2006). According to this design, repeated non-guided practice sessions given during a time of rapidly improving competence permits a high frequency of observations relative to the rate of change. Hence, changes in reasoning become visible at the very moment they happen, enabling the discovery of natural developmental pathways. These developmental pathways may be considered natural, since the practice sessions include no explicit forms of intervention, such as the provision of elaborate instructions or prompting (Flynn & Siegler, 2007; Siegler & Crowley, 1991; Siegler 2006). It is considered that the acquisition and development of cognitive abilities may show differing pathways when acquired through more 'natural' unprompted opportunities than when resulting from instruction. These potentially differing pathways make it useful to examine both in combination (Kuhn, 1995; Bjorklund, Miller, Coyle & Slawinsky, 1997; Opfer & Siegler, 2004; Tunteler et al., 2008). Therefore, in addition to unprompted repeated practice, we included instruction derived from a dynamic test.

Dynamic testing has become increasingly popular for the study of inductive reasoning (e.g., Bethge, Carlson, & Wiedl, 1982; Resing, 2000; Tzuriel, 2000; Tzuriel & Flor-Maduel, 2010; Tzuriel & Kaufman, 1999). Key to this approach is the incorporation of feedback and training during the testing phases (Sternberg & Grigorenko, 2002; Elliott, 2003; Swanson & Lussier, 2001). Conventional, static tests are considered to be means to assess already developed abilities. Dynamic modes of testing are designed to assess developing or yet-to-develop abilities which are the products of underlying, but often unrecognized, cognitive capacities (e.g., Hessels, 2000; Elliott, 2003; Lidz & Macrine, 2001; Resing, 2006; Sternberg et al., 2002; Sternberg & Grigorenko, 2006). Dynamic testing has been found to be a means to gain insight into cognitive and meta-cognitive strategies used by the examinees, their responsiveness to examiner assistance and support, and their ability to transfer learning from the test situation to subsequent unaided situations (Elliott, 2003). In this study, we examined the influence over time of a dynamic approach upon children's inter- and intra-individual developmental trajectories in analogical reasoning.

We also investigated the relationship of working-memory capacity to children's developmental trajectories. An accumulating body of evidence suggests that working-memory capacity is central to reasoning tasks such as the solving of analogies (Tunteler & Resing, 2010; Halford, Wilson & Philips, 2010; Morrison et al., 2004; Primi, 2001) and to learning in school (e.g., Alloway & Alloway, 2010; Alloway, Gathercole, Kirkwood, & Elliott, 2009; Holmes, Gathercole, & Dunning 2009; Swanson, 2008). In many studies the manner and extent to which inductive reasoning is related to working-memory capacity have been explored (e.g., Arendasy & Sommer, 2005; Meo, Roberts, & Marucci, 2007; Richland, Morrison, & Holyoak, 2006; Viskontas, Morrison, Holyoak, Hummel, & Knowlton, 2004; Waltz, Lau, Grewal, & Holyoak, 2000). When solving analogies, children's working-memory appears to be particularly important for encoding and processing the terms of the analogy (Sternberg & Rifkin, 1979).

Working-memory may be considered as the workspace for construction of relational representations for solving a given analogical task while using knowledge stored in semantic memory. This workspace is limited in the number of relations that can be processed in parallel although these typically increase with age and maturation. However, complex relations can be recoded into representations of lower complexity or be segmented into smaller parts in order to process them serially (Halford, et al., 2010; Halford, Wilson & Philips, 1998).

The type of relationship or task that needs to be managed appears to be influenced by a differential involvement of separate components of working-memory. Various components have been investigated in a variety of inductive reasoning or academic tasks (e.g., Raghubar, Barnes & Hecht, 2010; Alloway & Passolunghi, 2011). The age of the child and the differential involvement of these components in different types of tasks were first demonstrated by Alloway, Gathercole & Pickering (2006). In line with Baddeley and Hitch's (1974) working-memory model, they found that children as young as 4 years exhibit a structural organization of memory into a domain general component for processing information and verbal and visual-spatial domain specific components for storage. Furthermore, they found that these components could be assessed in a reliable way. In the present study we explicitly focused on the differential involvement of verbal and visual-spatial working-memory components, to examine their possible role in respect of analogical reasoning development in second graders. We thought it important to examine these components separately with a working-memory assessment that made sufficient storage and processing demands (Alloway, 2007) and which would help us explore their influence on analogical reasoning (Resing, Xenidou-Dervou, Steijn & Elliott, 2012).

Our type of data is traditionally analyzed by means of repeated measures analysis as it involves undertaking the same assessments at intervals over time for a given set of individuals. While repeated measures analysis does not enable the researcher to include in their analyses the variation between individual children's trajectories of performance, multilevel analysis – applied in a specific manner suited for longitudinal data - does enable the researcher to include in their analyses children's individual variation over time (e.g., Van der Leeden, 1998). Typically when employing multilevel analysis data, the individual participants are considered to be the first level units, and one or more grouping variables, for example, school or region, form the units for the higher level(s) within the model. Multilevel analysis of longitudinal datasets, on the other hand, allows one to analyze individual children's growth over time at a macro level, instead of at a micro level. Here the repeated measurements are viewed as the first level units, nested and correlated within individual children, who serve as the second level units (Hox, 2002, 2010; Kreft & De Leeuw, 2007; Snijders & Bosker, 1999; Van der Leeden, 1998). By modeling varying regression coefficients at the session level (Level-1), this form of multilevel analysis yields growth trajectories that typically vary for each individual child

(Level-2). Additionally, this form of multilevel analysis enables the inclusion of two types of explanatory variables in the model: time constant and time varying variables. As a result, it becomes possible to model both the average growth trajectories of each group, as well as the individual growth trajectories of each child (Hox, 2002). Thus analyzing our microgenetic data with this form of multilevel analysis enabled us to inspect growth trajectories (Level-1) for each individual (Level-2) and investigate systematic variation between these trajectories as a function of our background variables – the verbal and spatial working-memory components – and experimental treatment, the dynamic test (Van der Leeden, 1998).

In summary, the main focus of the current study was upon examining inter- and intra-individual developmental trajectories of analogical reasoning in a dynamic test and non-guided practice setting. This differed in several ways from earlier work (e.g., Primi, 2001; Hessels-Schlatter, 2002; Resing, Tunteler, De Jong, & Bosma, 2009; Tzuriel & George, 2009). Our explicit objective was to display by means of a relatively novel approach both children's individual growth trajectories in analogical reasoning performance (a) and systematic variation between these trajectories based on the experimental treatment and background variables – verbal and visual-spatial working-memory capacity (b) (Hox, 2002; Kreft & De Leeuw, 2007; Snijders & Bosker, 1999; Van der Leeden, 1998).

In the current study we, therefore, investigated analogical reasoning performance in second grade children by means of the microgenetic research method and multilevel analysis. In particular, we examined the relationships over time between repeated non-guided practice in analogical reasoning in isolation and repeated non-guided practice combined with a dynamic test session based on the 'graduated-prompts-technique' (e.g., Resing & Elliott, 2011) in children with differing levels of verbal and spatial working-memory capacity.

The objectives of the current study were to examine the inter- and intra-individual developmental patterns and rate of change in analogical reasoning between children, who (a) did or did not receive a dynamic test session, and (b) exhibited larger or smaller verbal and/or spatial working-memory capacity. With respect to (a) it was hypothesized that children who engaged in non-guided practice alone would increase their analogical reasoning performance over time, if they also exhibited greater working-memory performance. However, children who additionally received a dynamic test session were expected to show greater improvement over time, displaying the greatest rate of change after dynamic testing (e.g., Resing, 2000). With respect to (b) it was hypothesized that spatial working-memory capacity would be particularly important for analogical reasoning performance at the first session (e.g., Logie, Gilhooly, & Wynn, 1994; Rasmussen & Bisanz, 2005, Tunteler et al., 2008). In contrast, verbal working-memory capacity would be less influential (Alloway & Gathercole, 2006; Haavisto & Lehto, 2004; St. Claire-Thompson & Gathercole, 2006). Additionally, we expected that spatial working-memory capacity would influence improvement through repeated practice alone. Children in grade two with smaller spatial working-memory capacity were expected to display few changes in analogical reasoning through non-guided practice alone as their workspace for constructing relational representations is more limited (Halford et al., 2010). However, they were expected to exhibit a rather more rapid rate of change in analogical reasoning after dynamic testing (e.g., Carr & Schneider, 1991). The rationale was that dynamic testing was expected to alleviate any working-memory limitations by breaking down the analogical reasoning process into smaller steps that could be processed serially and by providing relational knowledge (Halford et al., 2010; Morrison, Doumas, & Richland, 2011). Children with larger spatial working-memory capacity, on the other hand, were expected to show a more gradual pattern and rate of change over time through repeated practice alone, while receiving additional benefit from the dynamic test (e.g., Tunteler & Resing, 2010).

3.2 Method

Participants
Participants were 32 children aged 7-8 years (18 boys; 14 girls) with a mean age of 90.1 months (*SD* = 4.7 months). They were selected based on their attendance in the second grade of two regular primary middle-class schools located in a midsize town in The Netherlands. Parental informed consent was obtained for each participant.

Design
During the first study weeks, each child's inductive reasoning and working-memory capacity were assessed by means of an Exclusion test and measures of spatial and verbal working-memory. Subsequently, a microgenetic two-pretest-two-posttest control-group design was employed (see Table 1). Children in the treatment condition received a dynamic test session while those in the control (practice) condition received a non-guided practice session. The non-guided practice session featured the same analogy tasks, but, as for the other practice sessions, children received no instruction, help or feedback. Non-guided practice sessions ranged from 20-40 minutes per child and were of equal duration for both conditions. The dynamic test session took 30-60 minutes per child for the treatment condition.

Table 1. Research design

	Session					
Condition	Pretest	1	2	D T[1]	3	4
Practice	x	x	X	-	x	x
DT[1]	x	x	X	x	x	x

Note: [1] DT = Dynamic Test; the practice-condition contained the same items as the DT-condition, but the practice-condition did not involve a dynamic test.

Instruments

Exclusion
Exclusion is a visual inductive reasoning subtest of a Dutch child intelligence test (RAKIT: Revisie Amsterdamse Kinder Intelligentie Test (Bleichrodt, Drenth, Zaal, & Resing, 1984)). The test consists of 40 items each comprising 4 geometric figures. Three of these figures can be grouped together on the basis of a rule that needs to be identified. The task requires the child to select the figure that, in each case, does not fulfill the rule.

Memory span–abstract

Memory Span-abstract is a subtest from the RAKIT that measures children's abstract memory span (RAKIT, Bleichrodt et al., 1984). The test consists of a booklet and small blocks both containing pictures of undefined shapes. The test items in the booklet have sequences of these shapes (2-7) that are shown for only 10 seconds to the child. Then, the child needs to reproduce these sequences with the blocks that have the same shapes printed on them. Although this test supposedly only measures memory span, it could also be considered to measure abstract-visual working-memory capacity. Simultaneous storage and processing are arguably involved in good task performance since the undefined shapes need to be manipulated into something more recognizable to remember, while being held in memory, in order to recall and reproduce longer sequences of these shapes. It is highly likely that our age-group draws on executive resources while performing this task (Alloway & Passolunghi, 2011).

Listening recall and spatial recall

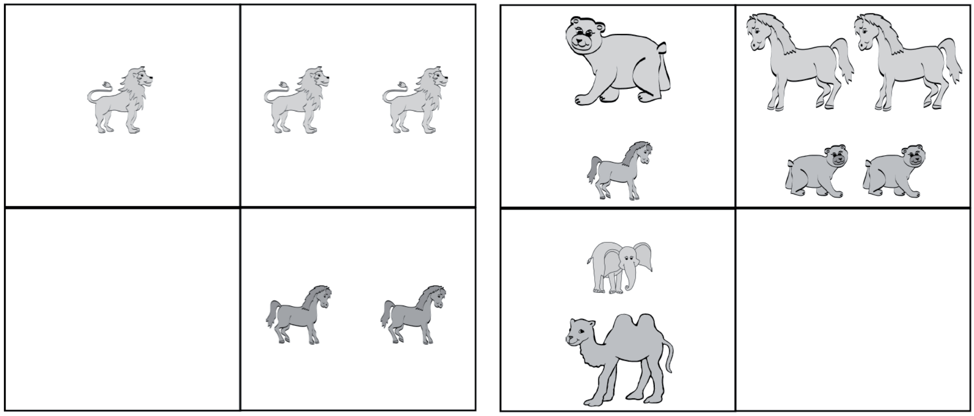
The screening measure from the computerized Automated Working Memory Assessment (AWMA) battery (Alloway, 2007) was used to measure verbal and visual spatial working-memory capacity. The AWMA measures involve both the simultaneous storage and processing of information. The listening recall task utilizes sequences of spoken sentences and the spatial recall task involves recalling the positions of arbitrary shapes that are rotated and/or flipped from left to right.

Figural analogies

The analogical reasoning task consisted of an age-adapted version of the concrete figural analogies measure developed by Stevenson and Resing (e.g., Stevenson, Resing, & Froma, 2009; Stevenson, Touw, & Resing 2011). Four parallel sets were created with 20 open-ended 2x2 figural matrix analogies (see Figure 1). In order to avoid responses based purely on visual recall, the parallel sets were designed to appear different by changing the animal-type and color of the figures of each item over sessions according to fixed rules. The figures consisted of various permutations of six types of animals with three familiar colors, and two sizes, features which would be easily recognized by the children concerned (Goswami, 1992). Items contained up to six transformations including, size, color, number, direction, position, and animal. Children's ongoing engagement was maximized by mixing the order of predicted difficulty of the items. This order of difficulty remained the same over sessions.

At the start of each session, the child was presented with a booklet containing the analogies and baskets with small animal cards for constructing the correct answers in accordance with the transformations used in the items. The examiner explained – while showing the animal cards – the features of the cards: three different colors of the same animal, a set of small and large cards for each animal and that the cards could be flipped. The examiner then turned to the first analogy and said that this was a 'kind of puzzle' with three boxes with animals and a fourth empty box (C-term or D-term), in which the child needed to construct the solution to 'the puzzle' using the animal-cards. After producing each solution, the child was asked how he or she had solved ´the puzzle´. Occasionally some children changed their solutions in response to their verbalizations. In such cases, the final physical arrangement of the cards was scored.

Figure 1. Examples of figural analogies used during non-guided and dynamic testing sessions (adopted from Stevenson et al., 2009)



Note: Left figure: the lion is yellow; the horse is red. Right figure: the small horse, small bears and camel are blue; the large bear, large horses and elephant are yellow.

Figure analogies training

The dynamic test material consisted of an age-adapted set of 7 concrete figural analogies similar to those employed in the other sessions (adapted from Stevenson et al., 2009; 2011). The steps involved are described in Appendix 3A. Unlike most other dynamic test formats, our measure proceeded from difficult to easy items. Where children needed assistance, we sought to provide the minimal amount of help required to solve the tasks independently, in accordance with Resing's (e.g., 1993, 1997) dynamic test format and the 'graduated-prompts-technique'.

The graduated prompts procedure used in the present investigation was originally pioneered by Campione, Brown, Ferrara, Jones & Steinberg (1985) and has been successfully utilized in several of our previous studies (e.g., Resing, 2000; Resing & Elliott, 2011; Resing et al., 2009, 2012; Resing, Steijn, Xenidou-Dervou, Stevenson, & Elliott, 2011). The procedure involves the use, during the dynamic testing session, of a series of adaptive and standardized, hierarchically ordered, metacognitive (self-regulating) and cognitive (task-specific) prompts that proceed from general (2 metacognitive prompts) to increasingly task specific (3 task-specific prompts). The prompts are only provided when and if a child is unable to proceed independently. This delivery of increasingly explicit prompts continues until the child reaches the correct solution. Children are provided with the minimum number of prompts possible to enable progression through the test. While our procedure contrasts with more traditional psychometric approaches whereby progression through the test typically moves from easier to harder items, we have found our approach valuable for enabling even the higher performing children to be trained from the outset, and for assisting all the children to use their newly learned strategies when solving easier items (e.g., Resing, 1993, 2000; Resing & Elliott, 2011; Resing et al., 2009, 2011, 2012).

Scoring
Working-memory capacity test-scores were converted into z-scores and subsequently into standard scores (*M*=100; *SD*=15). The two spatial memory tasks from the RAKIT (memory span-abstract) and the AWMA (spatial recall) were combined into a new variable: MemGrAVS. Verbal working-memory (i.e. the listening recall test) was labeled MemGrV. These two working-memory variables were each split into a 'lower score' and a 'higher score' category, based on the respective median scores on these variables of all 32 children. This yielded two equal groups of 'lower' and 'higher' scoring children for both working-memory variables separately.

The four analogical practice sessions were scored separately for each child. Every session included one booklet of 20 analogical matrices that was the same for all children. Each child received an 'Analogy Score' for each individual session that was the sum total of all analogies that were correctly solved during that session.

Analyses
Multilevel analysis (MLA) was used for analysis of the data. Traditionally, repeated measures analysis has been widely used to analyze data involving repeated measurements of the same individuals. However, microgenetic data sets can also be viewed as comprising multilevel data, where repeated measurements are nested within individuals (Hox, 2002, 2010; Kreft & De Leeuw, 2007; Snijders & Bosker, 1999; Van der Leeden, 1998). MLA appeared to be particularly valuable for the present study as it enabled us to inspect growth trajectories based on data obtained from repeated measurements (Level-1) for each individual (Level-2) and investigate systematic variation between these trajectories as a function of our background variables and experimental treatment (Van der Leeden, 1998). By modeling varying regression coefficients at the session level (Level-1), we obtained growth trajectories that were different for each individual child. Additionally, MLA allowed us to add two types of explanatory variables to the model: time constant and time varying variables. This allowed us to model both the average growth trajectories of each group, as well as the individual growth trajectories of each child (Hox, 2002, 2010).

For reference purposes, Appendix 3B displays the data structure and meaning of the variables used for the MLA. All of the variables contained a meaningful 0-point to facilitate interpretation (Hox, 2002).

3.3 Results

Before examining our research questions in detail, we checked for possible initial differences between children in the dynamic test and practice condition. The mean scores on Exclusion did not differ significantly, nor did the mean number of correct analogical solutions at session one. Means and standard deviations for 'Analogy Score' per session and condition are provided in Table 2.

Table 2. Means and standard deviations of analogy scores per session and condition

	Session				
	1	2	3	4	Total
Condition	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)	Mean (SD)
Practice (N = 16)	1.81 (3.17)	2.38 (3.59)	2.75 (4.85)	3.44 (5.32)	2.59 (4.32)
DT (N = 16)	3.00 (2.71)	4.31 (4.85)	8.38 (4.95)	8.06 (4.64)	5.94 (4.39)
Total (N = 32)	2.41 (2.95)	3.34 (4.27)	5.56 (4.90)	5.75 (4.99)	4.72 (4.35)

As described by Hox (2002, 2010), Multilevel Analysis for repeated measurement data was run with nine hypothesized nested models (see Table 3), to examine the inter- and intra-individual developmental patterns and rate of change in analogical reasoning between children, who (a) did or did not receive a dynamic test session in addition to repeated non-guided practice opportunities, and (b) exhibited a larger or smaller verbal and/or spatial working-memory capacity. As stated before, the nested models included the repeated measurements at level 1 and the individual children at level 2 (see Appendix 3B).

Table 3. Results of the likelihood ratio and AIC tests of the multilevel analysis for analogical solutions

	Model tests			
Model Progression[1]	Deviance	*λ(1)*	*P*	AIC
1. Intercept only (Null)	704.1			
2. + Session[2]	670.8*	33.3	< .001	678.8
3. + Spatial Working Memory	665.2*	5.6	.018	675.2
4. + Verbal Working Memory	665.0[4]	.2	.655	677.2
5. + Condition	640.3*	24.7	< .001	652.3
6. + Session Random[3]	624.3*	16.0	< .001	638.3
7. + Session*Condition	621.8[5]	2.5	.113	637.8[6]
8. + Session*Spatial Working Memory[7]	**618.8***	**5.5**	**.019**	**634.8**
9. + Condition*Spatial Working Memory	618.7	.1	.752	636.7

* Significantly better fit than former models at $p \leq .05$; [1]each successive model included one additional variable; [2]the time variable with 4 time points; [3]the slope of the time variable 'session' is modeled to vary across children in this and the following models; [4]non-significant in both the 'fixed effect only' and 'random slopes' model, and therefore no longer included after this point; [5]non-significant and therefore no longer included after this point; [6] the AIC diverts here from the likelihood ratio test [7]this interactions was the last one included in the final model, as the subsequent interaction did not improve the model any further.

Models progressed from those including only fixed effects to those with random slopes. Each successive model included an additional expected variable or interaction, after which it was compared to the previous model with a likelihood-ratio test to determine if the succeeding model had a significantly better fit than the previous one. For reference purposes, Table 3 also provides Akaike's Information Criterion (Akaike, 1987), although this was developed to compare non-nested models. Hox (2010), however, recommends the likelihood ratio test for nested models such as those used in the present study.

Rather than testing the hypotheses one by one in separate analyses, the best fitting model – according to the likelihood ratio test – was used to test our hypotheses by interpreting the interactions and the direct effects of the explanatory variables that made up the interactions together as an integrated system (Hox, 2002, 2010).

In relation to (a) it had been hypothesized that if children exhibited a greater spatial working-memory, repeated non-guided practice alone would improve their analogical reasoning performance over time, but children who had engaged in a dynamic test session would show greater improvement, with the greatest rate of change occurring after dynamic testing in relation to (b) it was hypothesized that spatial working-memory capacity would be particularly important for analogical performance at the first session and over time for gradual improvements in analogical performance through the non-guided practice sessions, but this would not prove to be similarly the case for verbal working-memory capacity.

After running the MLA, the eighth and final model (see Table 3) was proved to be the best fit. The likelihood ratio and the AIC yielded almost the same results. However, as stated above, the former was used to determine the final best fitting model, in accordance with Hox (2010). Regression lines are shown in Figure 2. For reference purposes, the regression equation for the best fitting model is displayed as Appendix 3C.

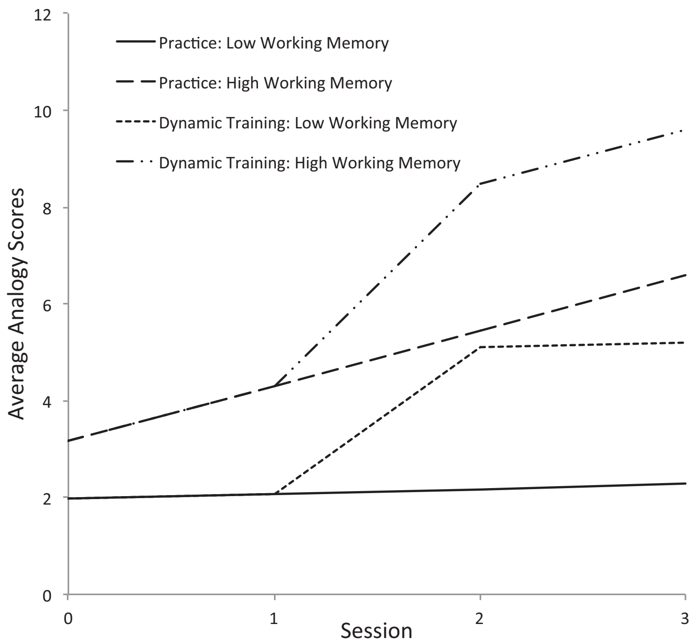


Figure 2. Regression lines per condition and working-memory group

Model 8 included two significant main effects: spatial working-memory and condition, and one interaction: session*spatial working-memory. These effects accounted for four 'subgroups' of children, each displaying a different rate of change over time (see Figure 2). The non-significant 'session' effect and the significant session*spatial working-memory interaction confirmed that children exhibiting a smaller spatial working-memory performance did not improve their analogical performance through repeated practice over time, but children exhibiting a larger spatial working-memory performance did improve their analogical performance through repeated practice over time. This improvement in analogical performance was, as expected, more gradual than the improvement that was induced by the dynamic test, regardless of children's smaller or large spatial working-memory performance. Also, the spatial working-memory main effect confirmed the influence of spatial working-memory on analogical performance at the first session. Verbal working-memory, as hypothesized, did not influence analogical performance. Furthermore, the non-significant interaction of session*condition showed that no significant losses or gains in dynamic-test induced analogical performance occurred at the fourth session for the dynamic test condition. Finally, the non-significant interaction of condition*spatial working-memory showed that no significant differences existed in dynamic test benefits between children exhibiting a smaller or larger spatial working-memory capacity. This confirmed that children with a smaller spatial working-memory capacity would be able to benefit from dynamic testing and improve their analogical performance in the same manner as their peers with a larger spatial working-memory capacity.

To further help interpret these results, we examined a graphical display of the individual children's growth trajectories (Hox, 2002). These trajectories, grouped on the basis of condition and spatial working-memory, are displayed in Figure 3. In general, children within the same condition and the same spatial working-memory group demonstrated similar growth trajectories. Nevertheless, their initial performance at session one displayed a fair amount of individual variability. This factor, in combination with spatial working-memory capacity, appeared to determine the growth trajectories for the sessions thereafter.

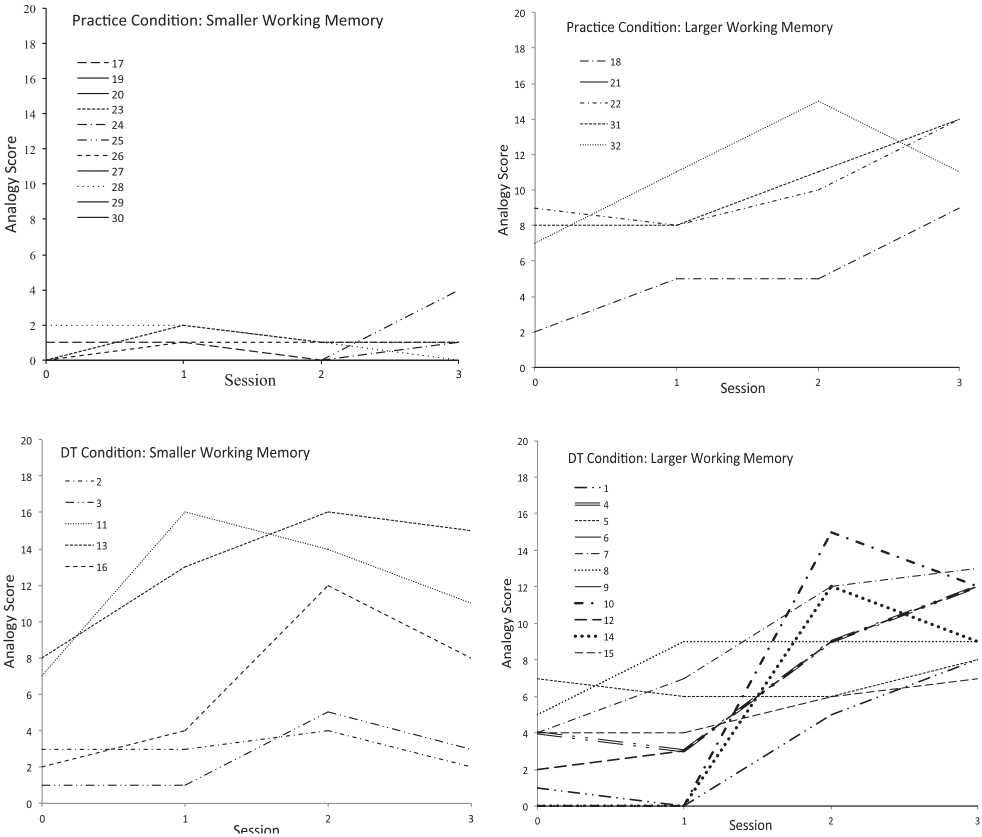
For the practice condition, individual growth trajectories of children exhibiting a smaller spatial working-memory capacity, demonstrated virtually no growth. However, several children exhibiting a larger spatial working-memory capacity displayed individual growth trajectories with a high initial performance and improved analogical scores over time induced by practice alone.

Children in the dynamic test condition with smaller working-memory capacity demonstrated one consistent pattern in common: their analogical reasoning performance deteriorated two sessions after the dynamic testing, although the MLA indicated that this reduction was not significant for the dynamic test condition as a whole. It is interesting to note that two children in this smaller spatial working-memory group obtained a rather puzzlingly high initial score.

Children in the dynamic test condition with larger working-memory capacity displayed the greatest variability in their initial analogical reasoning performance. Children in this group who obtained a low initial score displayed the fastest rate of change of all children across all of the groups; something that might be appropriately described as a 'light bulb effect'. In contrast, other children in this group, as expected, displayed a more gradual rate of change in analogical reasoning without displaying a drop in such performance at any time.

Finally, it is noteworthy that no individual child in any group obtained the maximum score of 20. In fact, only four children obtained a score of 15 or 16. This means that the most difficult items of the analogical task were too difficult for even the highest performing children. These implications of outcome are discussed below.

Figure 3. Individual growth curves for children in the practice condition with smaller working-memory capacity (upper left panel), children in the DT[1] condition with smaller working-memory capacity (lower left panel), children in the practice condition with larger working-memory capacity (upper right panel), and children in the DT condition with larger working-memory capacity (lower right panel).



Note: [1]DT = dynamic test. Caution: some individual growth trajectories are not (completely) visible due to overlap between children; children with solid lines portray complete overlap on the x-axis for all sessions (a 0-score for all sessions).

3.4 Discussion

This study's main aim was to examine the development of analogical reasoning in young children through the use of the microgenetic research method in combination with multilevel analysis. Specifically, these approaches were employed to investigate over time the inter- and intra-individual developmental trajectories and rate of change in analogical reasoning between children assigned to either a dynamic test or non-guided practice condition, while also considering verbal and abstract-visual-spatial working-memory capacity.

As hypothesized on the basis of group averages, children exhibiting larger spatial working-memory who received repeated non-guided practice alone improved their analogical reasoning performance over time, but not as much as those who also received dynamic testing. Furthermore, children who received dynamic testing subsequently displayed the greatest rate of change in performance. This confirms findings reported by Tunteler et al. (2008) and Tunteler and Resing (2010), who, in their microgenetic studies, found similar results with classical geometric analogies when children of grade 1 and 2 were provided with a short training procedure.

Second, as hypothesized, visual-spatial, but not verbal, working-memory capacity was related to a higher level of analogical reasoning performance at the first session. This finding supports earlier research in which differential involvement of working-memory components in various tasks was reported (Alloway & Passolunghi, 2011) and which demonstrated that such elements could be reliably measured in young children (Alloway et al., 2006).

As noted above, spatial working-memory was also related to improvement in analogical performance as a result of repeated practice. Dynamic testing added to this effect, with a subsequently greater average rate of change for all children, irrespective of their spatial working-memory capacity. This is in line with findings from an earlier study by Tunteler and Resing (2010). However, in their study, children with a smaller memory span were able to catch up with their peers after training. In our study there was no differential effect of the dynamic test for children with either small or large working-memory capacity. Therefore, those with greater spatial working-memory capacity continued to display superior analogical performance, even after dynamic testing.

It is unclear why this finding differs from that of Tunteler & Resing (2010). It is possible that the task in the present study was more demanding as it proved to be difficult for all of the participants, irrespective of their level of working-memory. Children in the two studies may have used their working-memory capacity in a different manner when solving the various problems (Halford et. Al, 2010). Interestingly, after inspecting children's individual growth trajectories, the results of the current study appear to have been caused in part by two children with a larger spatial working-memory capacity that increased their analogical reasoning performance very rapidly after dynamic training. This kind of performance change could, under different circumstances, be expected from children with smaller working-memory capacity solving a less complex type of analogical task.

A further explanation for the differing results comes from the observation of the individual growth trajectories of children with smaller spatial working-memory. Those in the dynamic test condition consistently displayed a drop in performance from the session immediately after the dynamic test to the next session. Such children in the practice condition, displayed no growth in analogical performance over time.

In contrast, those in the practice condition with larger spatial working-memory performed in line with our hypothesis that such children would gradually be able to increase their performance through practice alone. Yet, given that only four children were able to score as many as 15 or 16 out of 20, some of the analogies clearly proved too difficult to solve even given the assistance provided. As detailed information about individual children's growth trajectories was not provided in the Tunteler and Resing's (2010) study it is difficult to compare results across the two studies. A key strength of the present study, therefore, was the analysis and graphical display of children's individual growth trajectories over time.

The individual growth trajectories of children's analogical reasoning, against the backdrop of working-memory capacity, once again demonstrated that for the purposes of dynamic testing, the level of task difficulty should not be too high. Children can clearly profit from such a procedure, but only if the task to be solved lies within the individual's particular zone of proximal development (Vygotsky, 1978).

Our findings are also in line with Siegler's (2006) "overlapping waves" theory, which suggests that high initial variability of strategy use often predicts substantial later learning. Interestingly though, certain individual growth trajectories of children with larger spatial working-memory capacity, and who had received dynamic testing, suggested that a certain level of initial performance was not always necessary for rapid learning to occur. Dynamic testing appeared to have had a 'light bulb effect': after dynamic testing these children suddenly displayed a rate of change that exceeded all other learners.

The current study has produced findings that demonstrate that working-memory capacity is an important variable in the performance over time of subgroups that have already been identified as similar in earlier microgenetic studies (e.g., Tunteler et al., 2008; Tunteler & Resing, 2010). While this is consistent with some studies (e.g., Rasmussen & Bisanz, 2005; Morrison, et al., 2011; Alloway et al., 2009), it contradicts Resing et al. (2012), who found significant improvements in analogical reasoning performance irrespective of the working-memory level of the trained children. However, in this earlier study, working-memory was assessed in a simpler manner and a more traditional approach to data analysis was employed.

This distinction leads us to highlight two positive methodological aspects of the present study. Firstly, advanced working-memory tests were used to investigate working-memory components separately (Alloway et al., 2006). These tests might have resulted in improved assignment of children to working-memory groups. Secondly, the current study used a different means of analysis (multilevel analysis for repeated measurement data) that facilitated the inspection of individual growth trajectories in combination, rather than isolation, with systematic variation between these trajectories as a function of the background variables and experimental treatment (Van der Leeden, 1998).

Nevertheless, the current study was unable to display a clear and comprehensive picture of the underlying change mechanisms of the various subgroups. Similar to most studies with a microgenetic research design (Siegler, 2006), these only consisted of a few children per subgroup. This relatively small number of children did not permit us to arrive at comprehensive and strong conclusions. Possibly there was a lack of power to detect certain effects although, as we have shown, some were found in the study. Unfortunately, the small sample size prevented us from adding additional background variables, such as variability in analogical strategy use. Variable strategy use and children's subsequent learning appeared to be clearly present and this fits with findings from many microgenetic studies (Siegler, 2006) and with current theories about the relationship between working-memory capacity and analogical reasoning (e.g., Halford et al., 2010; Alloway et al., 2006).

Further studies with larger samples of children, as well as other age groups and ethnicities (e.g., Resing et al., 2009), larger training programs (e.g., Tzuriel & George, 2009), each employing long-term follow-up are needed to confirm the individual growth trajectories that we found. Additionally, more in-depth microgenetic research examining the combination of dynamic testing of analogical reasoning and working-memory assessment is likely to prove valuable.

Comparison of the variability and strategy use of the four subgroups identified in the current study could help to reveal specific strengths and weaknesses that influence particular learning trajectories. This information could be used to better predict children's learning trajectories and ameliorate potential problems by means of specialized support and instruction.

Dynamic testing may ultimately reveal particular forms of instruction, from metacognitive to more concrete (Resing, 2000), that are most powerful for children with different profiles. In addition, dynamic testing and working-memory assessment in combination may help to indicate the type of working-memory support or training most suited for an individual child (Morrison et al., 2011) although our current ability to offer classroom-based interventions for such difficulties remains sorely limited (Elliott, Gathercole, Alloway, Kirkwood, & Holmes, 2010).

Clearly, multiple sources of information are required to guide the design of high quality interventions for those with learning disabilities. It is contended that information from dynamic testing and assessment of working-memory capacity are likely to be valuable components of a holistic approach to maximizing children's learning. It is hoped that the present study has demonstrated the potential value of a unique approach that can aid the development of this goal.

Appendix 3A

Protocol for an example item of the dynamic test

Step	Instruction	Right Answer	Wrong Answer
0	0.1. Today we are going to make puzzles again. However, this time I will give you some help. 0.2. Just like the other times, there are animals in three boxes [*experimenter points to the boxes*], but there are no animals in the fourth box [*experimenter points to the empty box*]. 0.3. Again you may solve this puzzle by putting the animals in this empty box that you think belong there.	1. Yes, that's correct. 2. How did you solve the puzzle/ why did you put these animals here? [*experimenter continues to request information until the child gives no more information*] **Go to step 6**	1. Your solution isn't completely correct yet. 2. I will put the cards back and give you some help.
1	1.1. First, you think about where to start. 1.2. These boxes belong together [*experimenter points to the upper two terms of the analogy*]. 1.3. These boxes belong together in the same way [*experimenter points to the lower two terms of the analogy*] 1.4. These two boxes also belong together [*experimenter points to the correct two boxes*] 1.5. These two boxes [*experimenter points to the correct two boxes*] belong together in the same way. 6.1. What do you think should be put in the empty box?	1. Yes, that's correct. 2. How did you solve the puzzle?/ why did you put these animals here? [*experimenter continues to request information until the child gives no more information*] **Go to step 6**	1. It's not completely correct yet. 2. I will give some more help. [*experimenter puts the cards back*]
2	Try to solve the puzzle according to these steps: 1. Comparing the boxes 2. Thinking how the boxes belong together 3. Put down your answer with the cards 4. Check if your answer is correct by comparing the boxes.	1. Yes, that's correct. 2. How did you solve the puzzle?/ why did you put these animals here? [*experimenter continues to request information until the child gives no more information*] **Go to step 6**	1. It's (almost/ not completely) correct yet. 2. Let's look at it together [*experimenter puts the cards back*]

Step	Instruction	Right Answer	Wrong Answer
3	3.1. We start by comparing the boxes 3.2. How is this box changed to this box [A:B]? [*changes are requested, until the child provides no more information*] 3.3. And how is this box changed into that box [A:C]? [*changes are requested, until the child provides no more information*] 3.4. Now we shall think some more 3.5. These two [C:D] change like these two [A:B] 3.6. And these two [B:D] change like these two [A:C] 3.7. So.. how do we fill the empty box to solve the puzzle? 3.8. [If needed say: “You can put down the cards.”] 3.9. Let’s check, is this right? [*point to the row with boxes*] 3.10. Is this also right? [*point to the column with boxes*]	1. Yes, that’s correct. 2. How did you solve the puzzle?/ why did you put these animals here? [*experimenter continues to request information until the child provides no more information*] **Go to step 6**	OK, (it’s almost correct) I shall give you some more help. [*experimenter puts the cards back*]
4	**This box [A] changed to that box [B] because....** [*experimenter points to the changes and the boxes as he/she mentions them*] 4.1. the animals changed color. The dog is now red and the lion is now blue. 4.2. the animals changed size. The dog is now small and the lion is now large. 4.3. the animals changed in number. Of each animal there are now two. 4.4. the animals changed direction. The dog walked in this direction over here, but it walks in that direction over there. The lion walked in this direction over here and in that direction over there. 4.5. the animals changed places. The dogs are now at the top and the lions at the bottom. **This box [B] changed to that box [D] because....** 4.6. the animal changed into another animal. The lion became an elephant and the dog became a bear. 4.7. the animals changed size. The small lion is now a large elephant. The big dog is now a small bear. 4.8. the animals changed direction. The lion walked this way and the elephant now walks that way. The dog walked this way and the bear now walks that way.	1. Yes, that’s correct. 2. How did you solve the puzzle?/ why did you put these animals here? [*experimenter continues to request information until the child gives no more information*]	We are going to solve the puzzle together. [*experimenter puts the cards back*]

Step	Instruction	Right Answer	Wrong Answer
5	**5.1 We start with the animals. Which animals do we need?** [*experimenter points to the changes and the boxes as he/she mentions them*] *If answered incorrectly*: We need the elephants and the bear, because..... here [A:B] the animals remained the same. So here they also remain the same. Here [B:D] the lion changed to an elephant and the dog to a bear. So here you need elephants and bears. **5.2 Which color elephants do we need/ and which color bears?** *If answered incorrectly:* We need blue elephants and red bears, because… here [A:B] the animals changed color. So the elephants become blue and the bears red. Here [B:D] the elephant had the same color as the lion and the bear the same color as the dog. So we need blue elephants and red bears. **5.3 Do we need large or small elephant/ and large or small bears?** *If answered incorrectly*: We need small elephants and large bears, because… here [A:B] the animals changed size. So the elephants need to become small and the bears large. Here [B:D] the animal changed size. The small lion changed into a large elephant and the large dog changed into a small bear. So here the elephants need to be small and the bears large. **5.4 Do we need one or two elephants/ and one or two bears?** *If answered incorrectly*: We need two of each, because... here [A:B] the animals changed from one to two. So here we need two elephants and two bears. Here [B:D] the number of animals remained the same. So here the number also remains the same: two elephants and two bears. **5.5 In which direction do the elephants/ bears need to walk?** *If answered incorrectly*: The elephants walk this way [to the right] and the bears that way [to the left], because… here [A:B] the animals changed direction. So, the elephants now walk this way [to the right] and the bears walk that way [to the left]. Here [B:D] the animals changed direction. The lion walked this direction [to the right] and the elephant that direction [to the left], the dog walked this directions [to the left] and the bear that direction [to the right]. So here the elephants need to walk this direction [to the right] and the bears that direction [to the left]. **5.6 Do we place the elephants at the top or the bottom of the empty box/ and the bears?** *If answered incorrectly*: The elephants we place at the bottom and the bears at the top, because… here [A:B] the animals changed places. So the elephants need to be placed at the bottom and the bears at the top. Here [B:D] the animals remained in the same place. So the lions – that changed into elephants – we put at the bottom and the dogs – that changed into bears – we put at the top. [*The child may have already put the animals down and/or may change/ shift the animals at each sub-step. However, every sub- step needs to be mentioned*]		[In partnership with the child the right answer is created] 1. That is correct! 2. And why is this correct? [*experimenter continues to request information until the child gives no more information*]
6	**Give the correct explanation about the answer**		

Appendix 3B

Structure of analogical reasoning development data

Variable Names	Description	Range Min	Range Max
	Level-1		
Cons	Vector consisting of ones	1	1
Session	Test sessions: four measurement intervals	0	3
SessionEsq	Session squared	0	9
	Level-2		
Pupil_ID	Numbers assigned to individual pupils	0	32
Condition	Condition: 0 = practice; 1 = dynamic testing	0	1
MemGrV	Verbal memory group: 0 = low; 1 = high	0	1
MemGrAVS	Spatial memory group: 0 = low; 1 = high	0	1
	Dependent Variable		
Analogy Score	Score for the analogy test per child and session	0	16[1]

Note: [1]Although the maximum score possible is 20, no child at any session received a score higher than 16.

Appendix 3C

Regression equations for the final multilevel model

Regression Equation
Solutions Correct = 1.98 + .10 x session + 1.19 x spatial working-memory + 3.02 x condition + 1.04 x session*spatial working-memory.

Note: all variables contain a meaningful 0-point (including session). To obtain regression equations per subgroup, variables must be replaced with group codes and session numbers (practice condition & low working-memory = 00; practice condition & high working-memory = 01; training condition & low working-memory = 10; training condition & high working-memory = 11).