



<https://openaccess.leidenuniv.nl>

License: Article 25fa pilot End User Agreement

This publication is distributed under the terms of Article 25fa of the Dutch Copyright Act (Auteurswet) with explicit consent by the author. Dutch law entitles the maker of a short scientific work funded either wholly or partially by Dutch public funds to make that work publicly available for no consideration following a reasonable period of time after the work was first published, provided that clear reference is made to the source of the first publication of the work.

This publication is distributed under The Association of Universities in the Netherlands (VSNU) 'Article 25fa implementation' pilot project. In this pilot research outputs of researchers employed by Dutch Universities that comply with the legal requirements of Article 25fa of the Dutch Copyright Act are distributed online and free of cost or other barriers in institutional repositories. Research outputs are distributed six months after their first online publication in the original published version and with proper attribution to the source of the original publication.

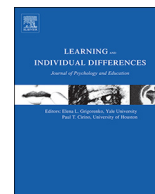
You are permitted to download and use the publication for personal purposes. All rights remain with the author(s) and/or copyrights owner(s) of this work. Any use of the publication other than authorised under this licence or copyright law is prohibited.

If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please contact the Library through email: OpenAccess@library.leidenuniv.nl

Article details

Stevenson C.E. & Hickendorff M. (2018), Learning to solve figural matrix analogies: The paths children take, *Learning and Individual Differences* 66: 16-28.

Doi: 10.1016/j.lindif.2018.04.010



Learning to solve figural matrix analogies: The paths children take

Claire E. Stevenson^{a,*}, Marian Hickendorff^b

^a University of Amsterdam, The Netherlands

^b Leiden University, The Netherlands



ARTICLE INFO

Keywords:

Analogical reasoning
Latent class analysis
Latent transition analysis
Hidden markov model
Relational reasoning

ABSTRACT

Analogical reasoning is essential for acquiring and integrating new knowledge and skills. Although much research has focused on this important skill, children's paths from non-analogical to analogical reasoning remain unclear. In this study, 388 children (ages 4–10 years) solved a series of figural analogies within a pretest-training-posttest design, with training comprising either multiple tries ($N = 196$) or tutoring feedback ($N = 192$). Working memory tasks were also administered. Latent transition analyses identified five phases with qualitative individual differences in children's analogy solving: duplication, idiosyncratic, beginner analogical, intermediate analogical and advanced analogical reasoning. Children's paths through these phases were not sequential; there was great variability between children and how they progressed through these phases. Working memory was related to children's reasoning phase at pretest, but not to their rate and path of change. Age and the type of feedback received during training were the clearest indicators of children's learning paths and rates of change.

1. Introduction

Analogical reasoning, the ability to recognize and use similarities and to generalize information from known situations to new ones, is essential for acquiring and integrating new knowledge and skills (Alexander, Jablansky, Singer, & Dumas, 2016). Children from a young age appear capable of reasoning by analogy (Goswami & Brown, 1990b; Rattermann & Gentner, 1998). However, factors such as lack of relational knowledge (Goswami, 1991), limited ability to shift focus from associative or perceptual features to relational features (Gentner, 1988; Hosenfeld, van der Maas, & van den Boom, 1997; Sternberg & Rifkin, 1979), and limited processing capacity (Halford, Wilson, & Phillips, 1998; Richland, Morrison, & Holyoak, 2006) form bottlenecks in its application. We know that children use a variety of approaches when learning to solve figural analogies and the most prominent approach prior to complete analogical reasoning is duplication, in which a copy of one of the problem elements is given as the solution (Siegler & Svetina, 2002; Tunteler, Pronk, & Resing, 2008). However, it is unclear *which* children transition *when* and *how* from being unable to correctly solve analogies to mastering analogical reasoning. The goal of this study was to shed light on individual learning phases and paths in the acquisition of figural analogical reasoning. In order to achieve this, we used dynamic testing, a microgenetic method of assessing cognitive ability and potential by including training in the assessment process, to observe and accelerate development (Elliott, Grigorenko, & Resing, 2010).

Analogical reasoning development is often examined with classical figural analogy tasks (see Fig. 1 for an example). The problem structure is usually described as A:B::C:?, where A-C represent subsequent problem elements (e.g., in Fig. 1, A is the figure with two small blue elements, B is the yellow elephant, C is the picture of two blue horses and the ? is the empty box that needs to be solved). Usually multiple-choice items are used (e.g., Chen et al., 2016; Siegler & Svetina, 2002). In the current study we used a constructed-response format which does not overly limit the children's solution choices and thereby provides a better picture of how they solve analogies (Stevenson, Heiser, & Resing, 2016). Constructed-response figural analogies can be solved by first *encoding* the objects on relevant dimensions (e.g., “two small blue elephants facing left”), then *integrating* the similarities and differences on these dimensions for each of the problem elements (e.g., “one of the elephants has gone away, it changed color, and looks the other way”), and finally *mapping* the pattern of changes in these dimensions (both horizontal and vertical) to mentally and physically construct the solution to the problem (Chen et al., 2016; Sternberg & Rifkin, 1979).

In order to successfully complete the first two steps in analogy solving, encoding and inference, children must be familiar with the problem elements and the relations between them (Goswami, 1989, 1991). Children as young as 4-years-old can solve figural analogies, choosing relational similarities above associative or perceptual similarities, when they are familiar with the causal relations in the problem domain (Goswami & Brown, 1990a). However, when solving figural

* Corresponding author at: University of Amsterdam, Psychological Methods, Nieuwe Achtergracht 166, 1018 WV Amsterdam, The Netherlands.
E-mail address: c.e.stevenson@uva.nl (C.E. Stevenson).

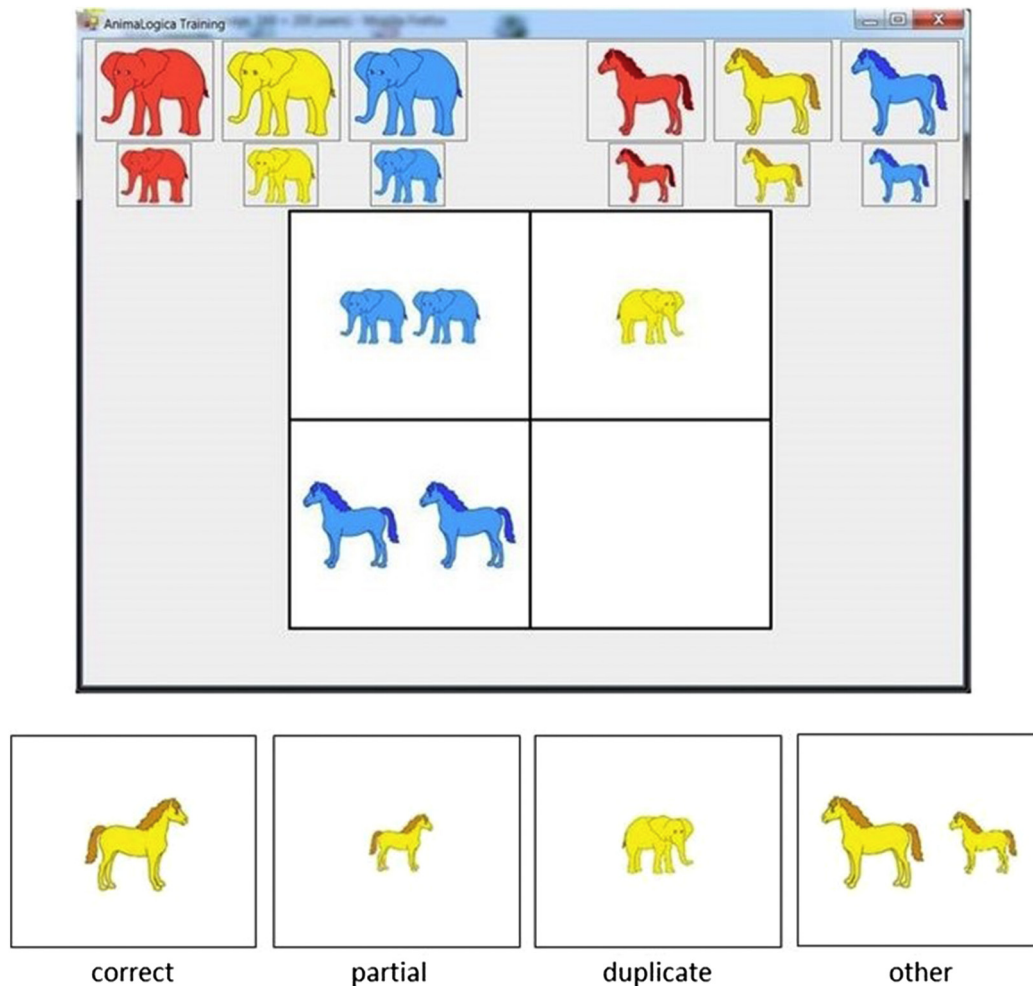


Fig. 1. A screenshot of a figural matrix analogy from AnimaLogica (Stevenson, Hickendorff, et al., 2013). Note that the solutions at the bottom are not multiple-choice options but, four examples of different types of solutions children provided.

analogies younger children generally choose duplicate solutions (i.e., copy of one of the matrix elements), a response based on perceptual rather than relational similarity. Siegler and Svetina (2002) attributed this to difficulty in encoding the relations required to solve the matrix. However, more recent studies in which children receive specific feedback on the relevant relations show that 5–6 year-olds are capable of encoding and representing these relations, and can become quite successful in solving figural analogies after just a few training sessions (Chen et al., 2016; Resing, Bakker, Pronk, & Elliott, 2017; Siegler & Svetina, 2002; Stevenson, Hickendorff, Resing, Heiser, & de Boeck, 2013; Tzuriel & Klein, 1985). Gentner (1988) proposed that a “relational shift” takes place in analogical reasoning where children’s representations of relations between analogy elements shift from simple feature comparisons (e.g., ‘This elephant is bigger than that one.’) to more abstract relational structures (e.g., ‘The elephants differ in size.’) and that this “relational shift”, rather than purely a question of age, steers the development of analogical reasoning (Gentner, Rattermann, Markman, & Kotovsky, 1995).

Improved executive functions, particularly working memory and inhibition control, is also considered a fulcrum for the transition from non-analogical to analogical reasoning in children (Halford et al., 1998; Morrison, Dumas, & Richland, 2011; Ropovik et al., 2016; Thibaut & French, 2016). Thibaut and French (2016) argue that young children have difficulty temporarily inhibiting the initial goal of the task ‘Find the figure that goes with C [in the same way as A goes to B]’ based on their research on the development of semantic analogy solving. This

may also explain why children solve figural analogies with duplicates of C, although this is dependent upon instruction. Richland et al. (2006) argue that young children have difficulty inhibiting perceptual similarities and therefore choose perceptual distractors above relational solutions. This account is less dependent upon instruction and could explain why children choose a duplicate in multiple-choice items. Less clear is how children respond to constructed-response items, where duplicates must be constructed rather than uninhibitedly chosen.

The last step in analogical reasoning, where relational similarities between two of the figures are mapped onto the remaining figure to construct the solution (i.e., the mapping and solving step), is considered very taxing for working memory as the mentally constructed solution must be retained in memory while selecting or constructing figures to answer the problem (Chen et al., 2016; Grossnickle, Dumas, Alexander, & Baggetta, 2016; Sternberg & Rifkin, 1979). The processing capacity perspective easily explains why many children solve figural analogies with a partially correct solution – where although most changes from the A to B term are correctly mapped to the solution, one or two are “forgotten”. Furthermore, a plethora of studies have shown that children’s working memory capacity and analogical reasoning ability are strongly related (e.g., Engel de Abreu, Conway, & Gathercole, 2010; Hornung, Brunner, Reuter, & Martin, 2011; Kail, 2007). However, an open question is whether working memory is related to children’s progression when learning to solve figural analogies, i.e., whether it can predict which children will transition from non-analogical strategies to full analogical reasoning. Some microgenetic, longitudinal and dynamic

testing studies show that working memory is linked to performance improvement from pretest to posttest (e.g., [Stevenson, 2017](#); [Stevenson, Heiser, & Resing, 2013](#)), whereas other studies indicate that it is not (e.g., [Resing et al., 2017](#); [Stevenson, Hickendorff, et al., 2013](#); [Tunteler et al., 2008](#)). Differences in working memory measures, age-group, sample size, design, training format or analysis technique do not explain these contradictory findings. Interestingly, there is little evidence that training working memory transfers to improvement in nonverbal problem solving, including analogical reasoning (e.g., [Melby-Lervåg & Hulme, 2013](#)), which weighs the evidence more in support of working memory not being the developmental mechanism driving the development of analogical reasoning.

The relational knowledge and processing capacity perspectives of analogical reasoning development combined provide many insights into the mechanisms of analogical reasoning development. Simulations with LISA, an influential computational model of analogical reasoning, imply that relational knowledge drives early development in analogical reasoning and that inhibition control and processing capacity form the bottleneck in later analogy solving ([Morrison et al., 2011](#)), and this hypothesis has received support from behavioral studies ([Richland & Burchinal, 2013](#)). However, neither perspective informs us about the phases and transitions that children go through when learning to solve analogies. In order to examine this further we turn to microgenetic and dynamic testing studies of children's analogical reasoning development; these have shown that children can learn to solve figural analogies spontaneously with practice ([Hosenfeld et al., 1997](#); [Tunteler et al., 2008](#); [Tunteler & Resing, 2002](#)) and that this can be accelerated with various forms of training ([Cheshire, Muldoon, Francis, Lewis, & Ball, 2007](#); [Siegler & Svetina, 2002](#); [Stevenson, Hickendorff, et al., 2013](#); [Tzuriel & George, 2009](#)). Although each study has shown large variation in the acquisition rate and path, learning to solve figural analogies appears to occur in at least three phases: (1) non-analogical reasoning characterized by duplicate solutions; (2) a transition phase in which diverse and inconsistent solutions are given, such as duplicates as well as partially correct answers; and (3) analogical reasoning with mostly correct solutions. We specify at least three phases because earlier work relied mostly on multiple-choice items with answer options including only duplication, partial and correct solutions, i.e. typical solutions from non-analogical, transitioning and analogical reasoning phases. For example, [Siegler and Svetina \(2002\)](#) found that 6-year-old children generally progressed from a phase with only duplicate solutions (i.e., copies of another problem element), to a phase with other errors (a combined category of partially correct solutions and perceptual distractors) and finally correct analogical solutions. Yet, similarly, [Hosenfeld et al. \(1997\)](#) examined 6- to 8-year-olds spontaneous analogy solving as it developed over eight sessions within a six-month period using open-format items and found signs of discontinuous change: a qualitative dip in performance with highly variable solution strategies when transitioning from non-analogical to analogical solutions was followed by a sudden jump in performance.

Practice and feedback in microgenetic and dynamic testing contexts clearly help children learn to solve analogies ([Chen et al., 2016](#); [Cheshire et al., 2007](#); [Resing et al., 2017](#); [Stevenson, Hickendorff, et al., 2013](#); [Tunteler et al., 2008](#); [Tzuriel & George, 2009](#)). Training with graduated prompts, a form of tutoring feedback that builds up from general to specific hints that guide the learner to the correct solution, appears especially effective in improving children's analogical reasoning ([Campione & Brown, 1987](#); [Resing & Elliott, 2011](#)) and helps children shift from non-analogical to analogical reasoning strategies ([Stevenson, 2017](#)). This may be due to learning to use relational terms, such as “color” or “size” rather than “red” or “big”, which is essential for the “relational shift” to occur ([Gentner et al., 1995](#)). However, the extensive graduated prompting sequence is not always necessary; multiple-try feedback, which does not provide feedback beyond correct/incorrect, appears nearly as effective as graduated prompting for

children who already apply analogical (relational) reasoning strategies ([Stevenson, 2017](#)). In the current study we used these two training approaches to accelerate children's learning of analogical reasoning and compared their effect on children's paths from non-analogical to analogical reasoning.

2. Current study

In sum, research shows that children transition from non-analogical to analogical solutions both spontaneously and with training, and in both cases there is much variability in the types of solutions they provide. However, it is unclear *which* children transition *when* and *how* from non-analogical to analogical reasoning. Age and the accompanying increases in relational knowledge clearly play a role, where younger children are less likely to correctly solve analogies. Working memory is also related to children's initial ability to solve figural analogies. But, training focusing on relational reasoning can help young children or children with lower processing capacity replace non-analogical (e.g., duplication) solution strategies with analogical reasoning so working memory may not be a necessary precondition for the shift from non-analogical to analogical reasoning. We used constructed-response figural analogies so that we could see which solutions the children actually constructed rather than selected based on limited response options that may not cover solutions from all phases of analogical reasoning development. The items contained familiar objects (i.e., animal figures) and relations (e.g., color, size) to ensure that the children had knowledge of the relations involved in the analogies. These were administered to 4–10 year-old children, an age range that was expected to cover all phases in analogical reasoning development that is thought to change most at 6–8 years of age ([Tunteler et al., 2008](#)), and within a dynamic testing context (pretest-training-training-posttest) to accelerate development. Our main research questions were: (1) Can we distinguish different analogical and non-analogical reasoning phases that reflect qualitative individual differences in children's matrix analogy solving?; (2) How do children transition between these phases when learning analogical reasoning?; and (3) To what extent are children's initial phases and transitions between phases related to age-group, working memory and type of training? We applied latent transition analysis (e.g., [Collins & Lanza, 2010](#)) on children's types of solutions on the pretest and posttest items to explore the children's phases and transitions in analogical reasoning and to examine whether age-group, working memory and/or type of training (graduated prompts or multiple-try feedback) were related to the individual differences herein. Based on the literature we hypothesized, first, that there are at least three main phases in analogical reasoning: non-analogical reasoning characterized by mostly duplicate solutions; a transition phase where a mix of duplicate, partially correct, and other solution types occur; analogical reasoning where duplicates are no longer given and analogies are consistently solved (partially) correctly. Second, we expected most children to transition sequentially through these phases, non-analogical followed by a transition phase and then followed by analogical reasoning as the literature reports few cases of children that return to duplication once children solve a set of figural analogies correctly above chance level ([Hosenfeld et al., 1997](#); [Siegler & Svetina, 2002](#)). Third, we hypothesized that the three discussed accounts (maturational, processing capacity, and relational knowledge) each contribute to the understanding of analogical reasoning development. Therefore we expected age-group (maturational and relational knowledge accounts) and working memory (processing account) to help predict a child's initial phase as well as transitions in analogy solving and that graduated prompts training (which provides relational terms during instructions) would lead to qualitatively greater gains (i.e., more transitions to an analogical reasoning phase) in analogical reasoning than multiple-try feedback.

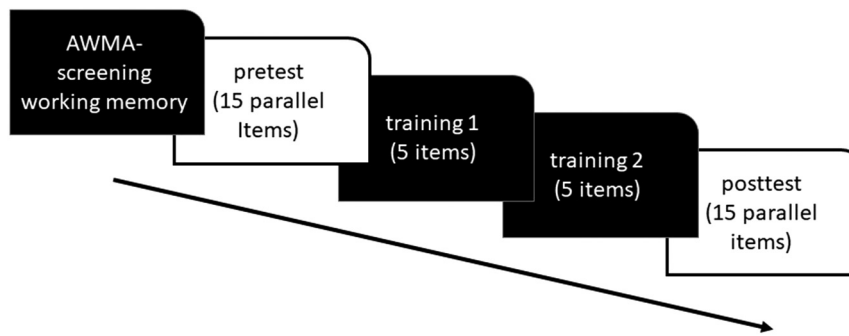


Fig. 2. Study design: There were five weekly sessions of approximately 30 min over the period of five weeks.

3. Methods

3.1. Participants

A total of 388 children from five age-groups (kindergarten through fourth grade, 66–90 children per age-group) were recruited from seven urban public elementary schools of similar middle class SES in the south-west of the Netherlands. The sample consisted of 186 boys and 202 girls, with a mean age of 7 years, 11 months (range 4–10 years). The schools were selected based upon their willingness to participate and written informed consent for children's participation was obtained from the parents or caregivers. The data collection for this study was approved by the Psychology Research Ethics Committee at Leiden University, the Netherlands.

3.2. Design & procedure

This study comprised a pretest-training-training-posttest design (see Fig. 2 for a complete depiction of study design). The children within each classroom were randomly allocated to one of two training conditions: (1) intelligent tutoring feedback or (2) multiple-try feedback. Four analogy dynamic testing sessions took place weekly and lasted 20–30 min each. Prior to the first analogy testing session the children were also administered the Automated Working Memory Assessment to assess verbal (subtest listening recall) and visuo-spatial (subtest spatial span) working memory (Alloway, 2007). All participants were tested individually in a quiet room at the child's school by educational psychology students trained in the procedure. The children received stickers after each session and a certificate upon completion.

3.3. Materials

3.3.1. Analogical reasoning assessment with AnimaLogica

AnimaLogica, a dynamic test of analogical reasoning including pretest, training and posttest phases, was used to test and train children in analogical reasoning (Stevenson, Hickendorff, et al., 2013). The figural analogies (A:B::C:?) comprise of 2×2 matrices with familiar animals as objects (see Fig. 1). The animals changed horizontally or vertically by color, orientation, position, size, quantity or animal type. The number of transformations – or object changes – provides an indication of item difficulty, where items with two transformations are considered the easiest and items with six transformations the most difficult (Mulholland, Pellegrino, & Glaser, 1980). The children were asked to construct the solution to the analogy using drag & drop functions to place animal figures into the empty box in the lower left or right quadrant of the matrix. A maximum of two different animals were present in each analogy. These were available in three colors (red, yellow, blue) and two sizes (large, small). The orientation (facing left or right) could be changed by clicking the animal figure. Quantity was specified by the number of animal figures placed in the empty box. Position was specified by location of the figure placed in the box – this

was automatically restricted to the top, middle or bottom location in the software.

The figural analogies could be solved horizontally or vertically. Essentially, the children selected an animal figure in a chosen color and size from the top of the screen. They then dragged and dropped the figure into the empty box. Here they use the mouse to drag it to a different position within the solution box or they could click on the figure to change the orientation. They could also easily remove an animal by dragging it to a position outside of the solution box. Each of the animal figures could be selected multiple times so the children could add as many figures as they thought were necessary to the solution, but there was a limit to how many figures could actually fit into the empty box as they were not allowed to overlap (the software automatically placed them side-by-side).

3.3.1.1. Solution coding. Each of the children's solutions were automatically coded into one of four categories. *Correct* solutions were those in which all features of the solution (i.e., animal, color, orientation, position, quantity and size) were correctly solved. Solutions were categorized as *partial* when one or two of the solutions features were incorrect (e.g., incorrect color and/or incorrect position). *Duplicate* solutions were a copy of the quadrant next to or above the solution box; if a solution could be categorized as either partial or duplicate then we coded this as a duplicate. *Other* solutions were those in which 50% or more of the solution features were incorrect (i.e., three or more transformations were incorrectly applied). See Fig. 1 for an example of each type of solution.

3.3.1.2. Example items. Before each testing or training session two example items were provided with the instruction “Here's a puzzle with animal pictures. The animals from this box have been taken away. Can you figure out which ones go in the empty box?”. If the child's solution was incorrect the correct solution was shown to assist task understanding, but without further explanation to avoid training the children further. During the testing phases the remaining items were administered without feedback.

3.3.1.3. Test items. The pretest and posttest items were isomorphs in which the items only differed in color and type of animal, but utilized the exact same transformations to ensure the same difficulty level. The number and difficulty level of the items differed somewhat per age group. The internal consistency of each of the versions was considered very good with Cronbach's $\alpha \geq 0.88$. In this study we focused on fifteen items that were included in each version of the test and solved by all children ($\alpha_{\text{pretest}} = 0.90$, $\alpha_{\text{posttest}} = 0.88$).

3.3.1.4. Training items. Each child solved 10 training items, none of which were present in the pre- or posttest to avoid memory effects. The items for each of the five age-groups were increasingly difficult, where the two easiest items were dropped and two slightly more difficult items were added for each subsequent age-group. Cronbach's alpha

Table 1
Overview of the prompts used in the graduated prompts condition.

Prompt	Verbal instruction
0	Here's a puzzle with animal pictures. The animals from this box have been taken away. Can you figure out which ones go in the empty box?
1	Do you remember what to do? Look carefully. Think hard. Now try to solve the puzzle.
2	This animal picture changes to this one. This one should change the same way.
3	So what changes here? Ok remember this one changes the same way.
4	See, this picture changes to this one because...
5	Which animal goes in the empty box? The elephant or the horse? What color should it be? Red, yellow or blue? ... Size? Quantity? Orientation? Position?

coefficient of internal consistency was computed per training item set across all first responses per item. The reliabilities ranged from $\alpha = 0.57$ to $\alpha = 0.79$; these are all considered satisfactory for 10 items.

3.3.1.5. Training procedures. The children in the *intelligent tutoring feedback condition* were trained according to the graduated prompts method, which has a long history within the context of dynamic testing of providing effective individualized short-term training of inductive reasoning abilities, such as analogical reasoning (Campion & Brown, 1987; Resing & Elliott, 2011). These consist of stepwise instructions beginning with general, metacognitive prompts, such as focusing attention, followed by cognitive hints, emphasizing the transformations and solution procedure, and ending with step-by-step scaffolds to solve the problem (see Table 1). In this version of AnimaLogica all instructions were provided by the trainer, and in some cases they were accompanied by visual effects supporting the explanations (see Fig. 3). A maximum of five prompts were administered. Once the children answered an item correctly they were asked to explain their answer to the trainer; no further prompts were provided and the next item was administered (for a more detailed description of the procedure see Stevenson, 2017). The children in the *multiple-try feedback condition* received auditory feedback during the training sessions on whether or not the outcome was correct. This was repeated until the item was solved correctly or five attempts were made to solve the item. After the fifth incorrect attempt the correct solution was shown before proceeding to the next item. If a correct solution was found before five attempts, then the next item was administered.

3.3.2. Automated working memory assessment (AWMA, Alloway, 2007)

The *Listening Recall* task was administered to measure verbal working memory. The child heard a sequence of spoken sentences (e.g., “bicycles can walk”) and was asked to say whether the sentence was true or false immediately following the sentence (e.g., “false”) and then to repeat the first word of each of the presented sentences (e.g., “bicycles”, ...). The sequence length began with a single sentence and increased by one sentence after four correct trials within a block of six trials. The task was terminated if less than four “first words” were correctly recalled. Scores were based on correctly recalled first words in the correct order. The total number of correctly recalled sequences was scored. The test-retest reliability of the Listening Recall subtest is 0.79 (Alloway, Gathercole, & Pickering, 2006).

The *Spatial Span* task was administered to examine visuo-spatial working memory. The child was presented with sequences of two shapes that were either facing the same (i.e., rotated but not mirrored) or opposite (i.e., rotated and mirrored) direction. Immediately after each presented pair the child was asked to convey whether the shapes were facing the same or the opposite direction. After each sequence of shape pairs, the child was asked to recall and point to the locations of the red dots that were displayed next to each right-hand shape. The sequences of shape pairs increased by one after four correct trials within a block of six trials. The task was terminated if less than four “red dot

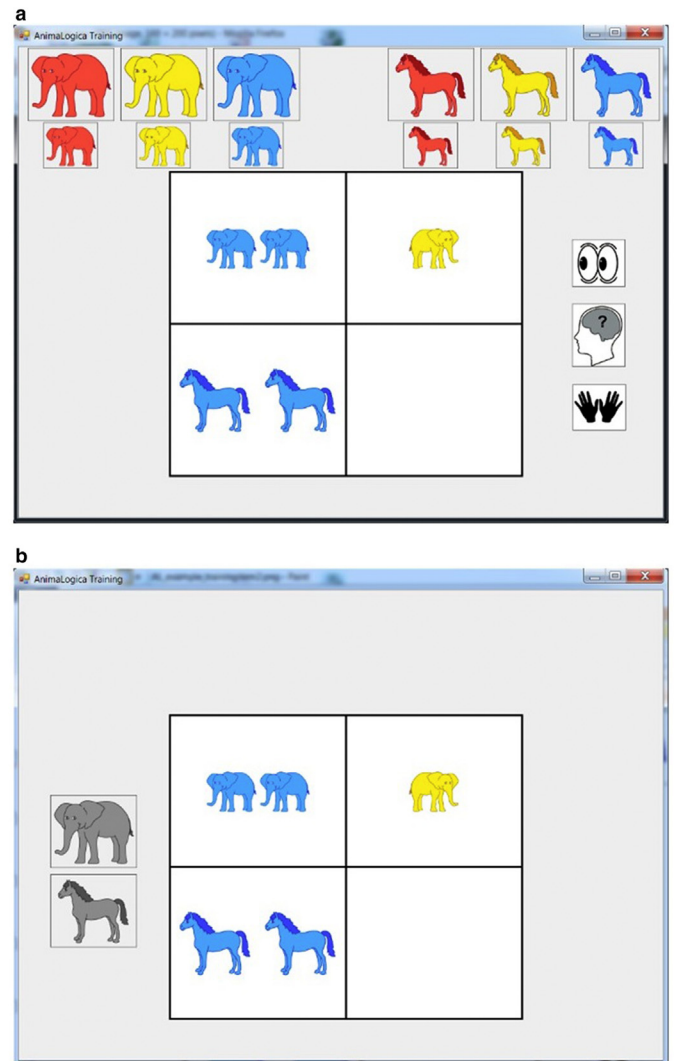


Fig. 3. (a) Depiction of visual effects emphasizes cues from prompt 1 to “Look carefully”, “Think hard” and then “Try to solve the puzzle” (these are not all shown at once). (b) Visual effects emphasizing prompt 5 where scaffolds are used to solve the puzzle: “Which animal belongs in the empty box?”.

locations” were correctly recalled. The total number of correctly recalled sequences was scored. The test-retest reliability of the Spatial Span subtest is 0.82 (Alloway et al., 2006).

3.4. Statistical analyses

The aims of the analyses were to: (1) identify phases reflecting qualitative individual differences in analogical reasoning, characterized by the types of solutions children used across the measurement occasions, (2) distinguish the paths children took from pretest to posttest while acquiring the skill of analogical reasoning, and (3) investigate individual differences in the occurrence of the phases and pathways related to the child variables working memory, grade, and training condition. We used Latent Class Transition models to achieve these aims with our complex categorical data (e.g., Collins & Lanza, 2010; Hickendorff, Edelsbrunner, McMullen, Schneider, & Trezise, 2017). *Latent class* models aim to identify qualitative individual differences between participants – “hidden” subgroups – based on a set of categorical response variables. Individuals within a subgroup share a characteristic response pattern that is distinct from the patterns in the other subgroup(s). These subgroups can represent different stages or phases in learning (e.g., Bouwmeester, Sijtsma, & Vermunt, 2004), such

as the different phases in learning to solve analogies that we aimed to identify. Furthermore, when data are collected on two (or more) measurement occasions latent class transition models can identify the developmental pathways. Discontinuous, phase-like change can be captured by the transition probabilities that represent the likelihood to move from one phase to any of the other phases. In the current study, we characterized the paths children take in learning analogical reasoning (i.e., between pretest and posttest) by analyzing whether and how the children moved between phases over time. Three sets of parameters are estimated: *initial class probabilities* reflecting the prevalence of each class at pretest; *conditional probabilities*, the probability of each of the solution categories per item per class which conveys the class-specific characteristic response patterns and is used to interpret the classes; the *transition probabilities*, which is the likelihood to move from any one class at pretest to each class at posttest.¹

The latent transition analyses (LTA) were conducted with Latent Gold 5.0 (Vermunt & Magidson, 2013a). In order to soundly capture the change during dynamic testing, we only used the items that were administered in all sessions in our analyses. Therefore, input data were the solutions on the 15 items administered at both pretest and posttest (no training items as these differed from test items), which were classified into one of four categories: duplication, partial analogical reasoning, correct, and other. The model fit procedure we used consisted of several steps. First, we ran LTA models (without covariates) with an increasing number of classes to find the optimal number of latent classes or phases. The decision to select a particular model was based on a combination of model-fit statistics, where we used the BIC as it has been shown to perform best (Nylund, Asparouhov, & Muthén, 2008), and conceptual appeal of the results (Collins & Lanza, 2010). Second, we checked to what extent the time-invariance of the classes was supported by conducting latent class analyses (LCAs) on the data for the pretest and posttest separately.

Third, after having selected the number of classes and checked the similarity of the classes resulting from the LTA (pretest and posttest simultaneously) and the separate LCAs (pretest and posttest separately), the child variables grade (sample-standardized), working memory (verbal and visual composite score, sample-standardized), training type (tutoring or multiple-try feedback), and training intensity (sample-standardized number of hints received during training) were added as covariates to the LTA. We tested the effects of the covariates for statistical significance with a Likelihood Ratio Test. This test statistic is based on the improvement in model fit (log-likelihood) when comparing the more complex model that includes the effect of interest compared to the simpler model without this effect. This yields a test statistic that can be tested against the chi-square distribution, with degrees of freedom equal to the number of parameters estimated for the effect of interest.

With a forward selection procedure we determined which covariate effects on (a) the initial class probabilities, and (b) the transition probabilities should be in the model. In modeling the covariate effects on transition probabilities, we used the “standard logits” option rather than the “transition logits” option, meaning that the effects of a covariate on the probabilities to move to each of the posttest classes did not vary across the pretest classes (Vermunt & Magidson, 2013b). In other words, the covariates affected the likelihood to move to each of the classes at posttest irrespective of the class children were in at pretest. This is much more parsimonious than estimating covariate effects on all transition probabilities, and the model fit turned out to be not significantly worse. To check whether all effects selected in the forward selection procedure contributed to the model, we checked whether

covariates with non-significant Wald-statistics (Vermunt & Magidson, 2013b) could be removed without significantly affecting the model fit.

Since covariates may affect the model parameters (conditional probabilities) and hence the interpretation of the model (e.g., Collins & Lanza, 2010), as a fourth step we checked whether the same number of classes would have been selected with the selected covariate effects (as determined by the forward selection and backwards elimination procedures) in the model. Therefore, we re-ran LTA-models with an increasing number of classes, but now with the covariate effects included, and selected the optimal number of classes based on model-fit statistics. The alternative to include the covariates already in the model selection procedure was not practically feasible, because there were too many possible effects (4 covariates, on 5 initial class size parameters and on 25 transition probabilities parameters) to include all of them.

After these four steps converged in the selection of one final model, the parameter estimates of this model were interpreted to answer the research questions. To address the first research questions concerning the identification of phases of analogical reasoning learning cross-sectionally, we interpreted the class-specific probabilities to use each solution category per item. To address the second research question concerning the characterization of children’s movement between phases over time, we interpreted the transition probabilities. To address the third research question, concerning the effects of child variables (grade and working memory) and training characteristics (training type and training intensity), we interpreted the selected effects of these variables as covariates on either (a) the likelihood to be in any one of the phases at pretest, and (b) the likelihood to move from any phase at pretest to any phase at posttest (the pathways).

4. Results

4.1. Descriptive statistics

Table 2 presents the descriptive statistics of the children’s analogy solving performance at pretest and posttest, training intensity as well as working memory composite scores, by grade and training type. Performance improved with grade on all four measures (note inverse of training intensity implies improvement), $F(4,383) > 54.2$, $ps < 0.001$. There were no differences between the two training conditions on pretest, posttest or working memory composite scores $F(1,386) < 1.6$, $ps > 0.05$. However, training intensity differed between the multiple-try and tutoring feedback conditions, with greater feedback requirements in the multiple-try group ($t(387) = 3.38$, $p < 0.001$). Analogy solving performance at posttest was higher than at pretest ($t(387) = -19.03$, $p < 0.001$). In addition, the four measures were each strongly related to the others ($ps < 0.001$).

4.2. Model selection procedure

In the first step, we estimated a set of latent class transition models with one to seven latent classes (see Table 3 for model fit statistics and supplementary materials² for details of all models). The model with five latent classes had the lowest BIC-value. Models with three to six classes were interpreted for conceptual appeal. The classes in the models with three, four and six classes could be mapped conceptually on those of the five-class model. In the three-class solution the first two (more advanced reasoners) classes were merged into one class and this was also the case for the middle two classes (transitioning); the last class (duplicators) was nearly the same as that of the five-class solution. For four-class solution, the middle two classes were redistributed into the middle three classes in the five-class solution, whereas the first and last class remained virtually identical. Similarly, the first three classes in the

¹ Because children were nested in classrooms, we calculated the intraclass correlation (ICC) of pretest correct/incorrect scored data to check for classroom differences. The ICC of 6% and design effect of 1.86 did not necessitate accounting for the hierarchical structure of the data in the analyses.

² The final data set and analysis results are available in the model-selection folder of our online supplementary material: <https://osf.io/8xn5r/files/>.

Table 2

Descriptive statistics of pretest and posttest performance (percentage correctly solved analogies), working memory composite score, training intensity (number of times feedback was given during training): means (SDs) and correlations.

	Grade					Training		Total	Correlations		
	KG	1	2	3	4	Multiple-try	Tutoring feed-back		Pre-test	Post-test	WM
Pretest	4.8 (10.3)	11.9 (18.4)	31.5 (26.1)	47.8 (23.8)	56.0 (19.1)	30.5 (29.0)	27.4 (26.9)	29.0 (28.0)	–		
Posttest	15.5 (16.8)	38.7 (26.4)	56.6 (20.4)	70.7 (20.2)	72.2 (16.4)	47.9 (30.9)	51.6 (27.4)	49.7 (27.2)	0.72*	–	
Working memory	15.4 (7.8)	22.1 (7.6)	27.4 (6.3)	33.1 (6.7)	34.1 (8.3)	25.7 (10.2)	26.4 (9.8)	26.0 (10.0)	0.63*	0.64*	–
Training intensity	27.8 (12.8)	19.9 (13.3)	11.9 (10.0)	9.5 (10.5)	8.7 (7.3)	18.0 (14.1)	13.6 (11.7)	15.8 (13.2)	–0.67*	–0.76*	–0.57*
n	75	90	87	66	70	196	192	388			

* $p < 0.001$.

Table 3

Model fit statistics of the estimated and compared Latent Class Transition models. The lowest BIC-value, a criterion to select the best model, is printed in boldface.

# Classes	LL	BIC	AIC3	# Parameters	df	Classification error
1	–13,947.6	28,163.4	28,030.1	45	343	0
2	–11,049.9	22,654.1	22,378.8	93	295	0.007
3	–10,359.9	21,572.3	21,148.9	143	245	0.019
4	–9781.8	20,726.0	20,148.6	195	193	0.030
5	–9521.2	20,526.8	19,789.5	249	139	0.042
6	–9387.9	20,593.9	19,687.2	305	83	0.047
7	–9264.6	20,693.1	19,618.3	363	25	0.059

Table 4

Model fit statistics of latent transition models with covariates.

Model	Effects on pretest phase occurrence	Effects on transition probabilities	LL	P ^a	BIC	LRT (df) ^b	p
M0	–	–	–9521.2	249	20,526.8		
Block A: predictors affecting pretest phase occurrence							
A1	Grade	–	–9376.0	253	20,260.1	290.5 (4)	M0 < 0.001
A2	Grade + WM ^c	–	–9353.0	257	20,238.1	45.9 (4)	A1 < 0.001
Block B: predictors affecting transitioning probabilities							
B1	Grade + WM	Grade	–9245.1	261	20,046.1	215.8 (4)	A2 < 0.001
B2	Grade + WM	Grade + WM	–9239.4	265	20,058.4	11.5 (4)	B1 0.021
B3	Grade + WM	Grade + WM + training condition	–9234.2	269	20,071.9	10.3 (4)	B2 0.035
B4	Grade + WM	Grade + WM + training condition + training intensity	–9188.7	273	20,004.8	90.9 (4)	B3 < 0.001
B5	Grade + WM	Grade + WM + training condition * training intensity	–9187.7	277	20,026.5	2.1 (4)	B4 0.71
B6	Grade + WM	Grade + training intensity	–9189.3	265	19,958.3	1.13 (8)	B3 0.99

^a P = number of parameters estimated.

^b Likelihood Ratio Test (LRT) involves a comparison between the current model including the predictor variable to be tested and the previous model without that predictor variable.

^c WM = working memory.

five-class solution were redistributed into the first four classes in the six-class solution, whereas the last two classes remained the same. The three-class and four-class models were deemed a bit too crude by not distinguishing between children with idiosyncratic solutions from those who were already in one of the more developed phases, whereas extra differentiation resulting from the six-class model was not very relevant conceptually. Based on conceptual appeal and the relative fit statistics, we therefore selected the five-class model.

In the second step, separate latent class analyses for the pretest data and posttest data were conducted to check the time-invariance of the latent classes resulting from the latent transition analysis. For the pretest data a four-class solution was optimal. The patterns of conditional probabilities mirrored those of the first four classes in the latent transition analyses, as did the class sizes. Since the fifth class was nearly empty at pretest in the LTA it was to be expected that this class did not emerge as a separate class in the LCA on the pretest data. These patterns corroborate with those of the latent transition analysis described in Table 4. The latent class analysis on the posttest data pointed to the four-class model as the optimal one, but the interpretation of the five-class solution yielded very similar class interpretations and class sizes as the latent transition analysis did. Taken together, the latent class

analyses on pretest and posttest data separately yielded very similar solutions to the latent transition analysis on pretest and posttest simultaneously, supporting the time-invariance of the classes.³

Third, we determined which covariate effects should be included. There were four covariates: the child variables grade and working memory, and the training variables training type (multiple-try feedback versus tutoring feedback) and training intensity. These covariates could be related to phase prevalence and/or transition probabilities. We examined this in two blocks. In block A we determined whether the two child variables grade and working memory had a significant effect on pretest-occurrence of the five phases of analogical reasoning (i.e., the initial class sizes). In block B, we determined whether the child variables grade and working memory and the training variables training type and training intensity, had a significant effect on the pathways between phases (i.e., the transition probabilities).

Table 4 shows model fit statistics and Likelihood Ratio Tests (LRT) testing the different predictors for significance. The results of the models fitted in block A show that grade and working memory

³ The output and a comparison of these analyses can be found in the time-invariance folder of our online supplementary material: <https://osf.io/8xn5r/files/>.

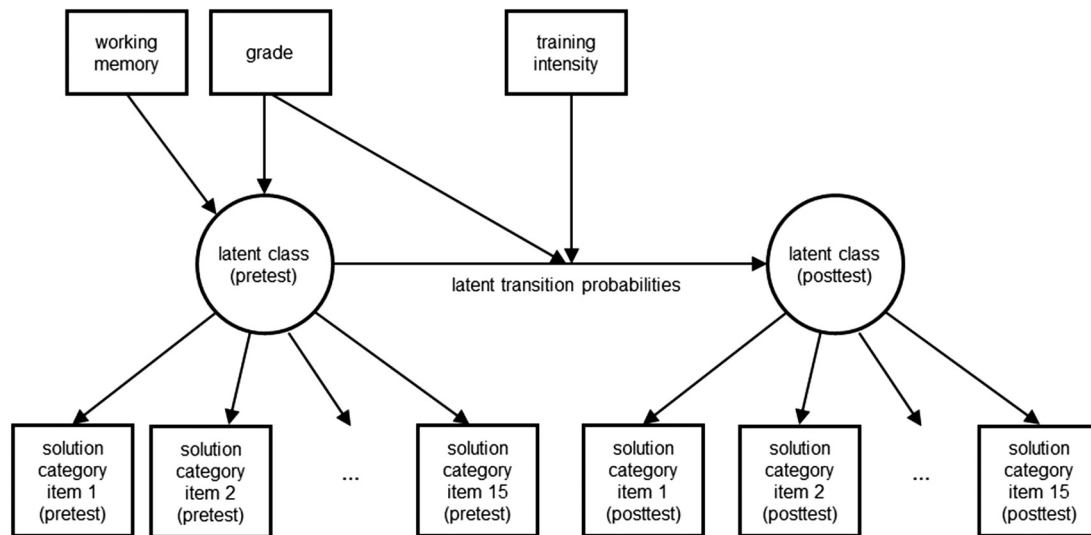


Fig. 4. Graphical representation of effects in final model B6.

measures had a significant effect on pretest phase occurrence, and models B1 and B2 showed that these variables also affected the transition probabilities. Further adding training condition (model B3) and training intensity (model B4) were both significant, whereas the interaction effect between training condition and training intensity (model B5) was not. Finally, we checked the Wald statistics of each of the effects in model B4. The effects of working memory and training condition on transition probabilities were not significant (Wald < 0.59). Removing these two effects in model B6 did not result in a significant decrease in model fit and this model had the lowest BIC-value across all models. Therefore, we selected model B6 as the final model. In this model grade affected both pretest phase occurrence and transition probabilities, working memory only affected the pretest phase occurrence, and the training intensity affected the transition probabilities.

Since adding covariates may affect the model solution, we re-estimated models with one to seven latent classes but now with the covariate effects as in model B6. Again the model with five latent classes had the lowest BIC-value. Second, we checked the congruence of the five-class model without covariates (model M0) and with covariates (model B6), and the interpretations were highly comparable.

In conclusion, the model selection procedure resulted in the final model with five latent classes, where the children's grade and working memory affected the pretest phase occurrence, and the children's grade and intensity of the training (between pretest and posttest) affected the transition probabilities, as illustrated in Fig. 4. This model had a classification error of 0.033 and an R-square entropy of 0.946, meaning that the quality of the classification was good. The model parameter estimates of this final model are interpreted to address the three research questions.

4.3. Phases in analogical reasoning

The first research aim was to identify phases reflecting qualitative individual differences in analogical reasoning, characterized by the types of solutions the children used across pretest and posttest. Fig. 5 depicts the conditional probabilities for each of the solution categories in each latent class – which we call phase from now on. To ease interpretation, we averaged these probabilities across items of similar difficulty (easy items were solved correctly by 50% or more of children, intermediate items by 35 to 45%, and difficult items by < 35%; there were 5 items in each category). Now we provide a conceptual interpretation of the five latent phases in the best fitting model. Please note that the order of the latent phases from the analysis was arbitrary (label

switching problem, Collins & Lanza, 2010), so we report these in the order we consider most logical.

Latent class 1 was labeled *duplication phase* because in this phase children were very likely to use duplication on all items. Latent class 2 was labeled *idiosyncratic phase* because these children solved easy items in a variety of ways, although partial analogical solutions were most common, and they generally solved more difficult items with other, idiosyncratic, non-analogical reasoning strategies. Children in latent class 3, labeled *beginner analogical phase*, were likely to solve easy items correctly and provide partial analogical solutions on average and more difficult items. Children in latent class 4 generally solved easy items correctly and had roughly a 50–50% chance of solving more difficult items correctly or partially correct; therefore this class was labeled *intermediate analogical phase*. Finally, children in latent class 5 solved most items correctly, and when mistakes were made the solution was partially correct; thus, we labeled this class the *advanced analogical phase*.

In addition to qualitative inter-individual differences the five latent phases identified, an interesting pattern of intra-individual differences within the phases also emerged. Specifically, for children in the idiosyncratic, beginner analogical, and partial analogical reasoning phases the likelihood of more advanced solutions (correct, partial) decreased at the cost of less advanced solutions (duplication, other) for more difficult items. Interestingly, this pattern did not emerge in the duplication phase, where children consistently applied duplication irrespective of item difficulty.

4.4. Paths from pretest to posttest

Our second research aim concerned possible learning paths children take when learning analogical reasoning. To examine this we focused on the probabilities that a child would move from a particular phase at pretest to each of the (other) phase at posttest and the proportion of children in each class at pretest and posttest. Table 5 presents all probabilities and Fig. 6 illustrates them graphically.

We see that children were much more likely to move to a more advanced phase from pretest to posttest than to stay in the same phase or move to a less advanced one (Fig. 6). First, children were unlikely to stay in the duplication phase (only 13% of the pretest-duplicators did) and instead tended to transition to beginner or intermediate analogical reasoning. Second, it was very unlikely to transition to the duplication phase from any other phase. Consequently, the duplication phase was much more prevalent at pretest (24%) than at posttest (6%). Third, children who started off in the idiosyncratic, beginner analogical, or

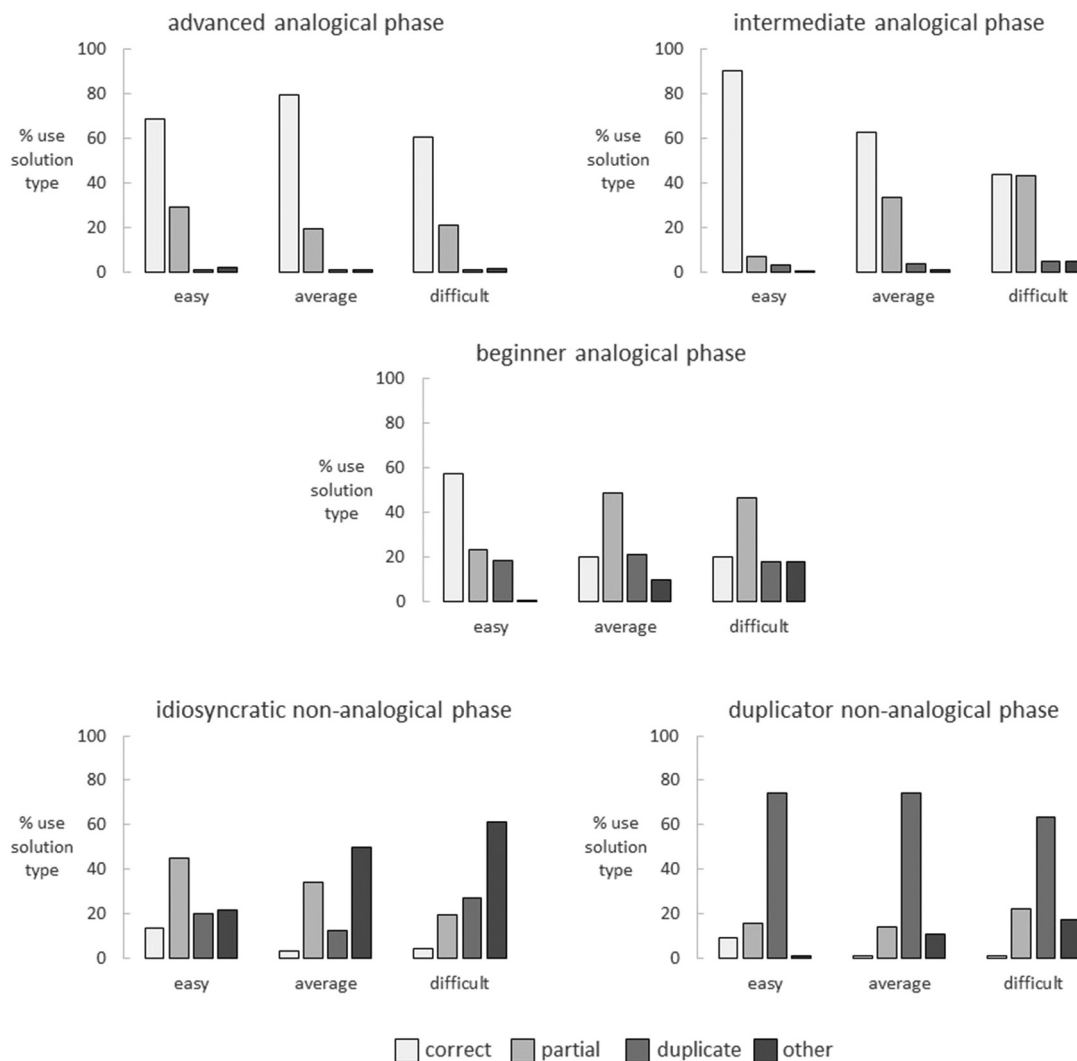


Fig. 5. Distributions of solution types (correct, partial, duplicate, other) within each phase of analogical reasoning for easy (overall proportion correct ≥ 0.5), average (proportion correct ≥ 0.3 and < 0.5) and difficult items (proportion correct < 0.3).

intermediate analogical phases were very likely to move “up” one or two phases. Fourth, at pretest the advanced analogical phase was practically empty ($< 0.1\%$ of the children) whereas 36% of the children were in that phase at posttest. Our results thus far indicate that most children’s analogical reasoning improved from pretest to posttest – progressing towards more advanced analogical reasoning phases.

4.5. Individual differences predicting analogical reasoning phases and learning paths

To address the third research question we interpret the results of the covariate effects: grade and working memory affecting pretest phase

Table 5
Marginal proportion of occurrence of each latent class at pretest and posttest, and transition probabilities.

Phase	Pretest size	Probability to transition to					Posttest size
		1. Duplication	2. Idiosyncratic	3. Beginner	4. Intermediate	5. Advanced	
1. Duplication	0.24	0.13	0.14	0.34	0.38	0.01	0.06
2. Idiosyncratic	0.22	0.13	0.28	0.28	0.28	0.03	0.12
3. Beginner	0.19	0.00	0.03	0.20	0.38	0.38	0.19
4. Intermediate	0.34	0.01	0.01	0.01	0.18	0.77	0.27
5. Advanced	0.00	0.02	0.02	0.02	0.46	0.48	0.36

occurrence, and grade and training intensity affecting the transition probabilities. The multinomial regression coefficients are in Table 6.

The effects of grade and working memory on the phase occurrence (top rows of Table 6) at pretest are shown in Fig. 7. These two plots show a very similar pattern: older children (i.e., upper grades) and children with greater working memory are more likely to be in the intermediate analogical phase and less likely to be in the duplication or idiosyncratic reasoning phases; the likelihood of being in the beginner analogical reasoning phase peaks around the second grade (7–8 years old) and for children with average working memory scores.

The effect of grade and training intensity on the transition probabilities (bottom row of Table 6) are in opposite directions. That is,

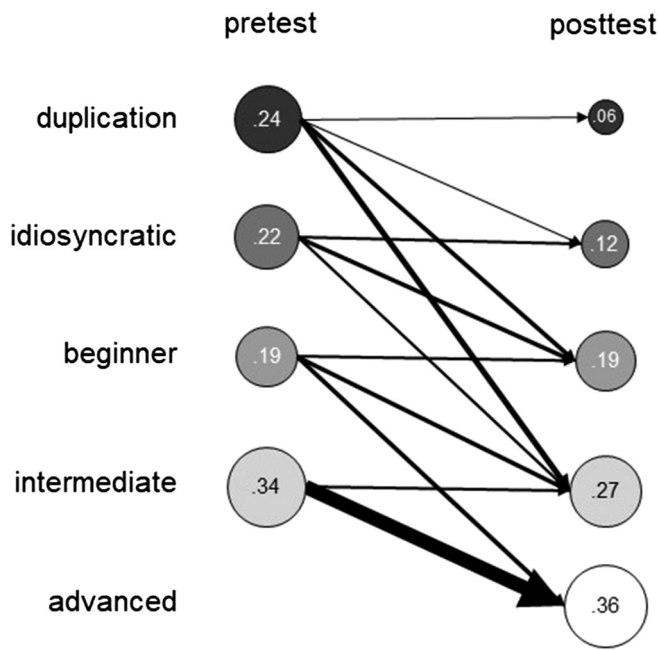


Fig. 6. Graph of children's learning paths from pretest to posttest to/from each phase: non-analogical duplication, non-analogical idiosyncratic, beginner analogical, intermediate analogical, and advanced analogical reasoning. Node sizes are proportional to number of children in that phase; darker, thicker lines represent more children transitioning from the pretest to posttest phase. Only paths taken by at least 3% of all children are displayed.

when children get into higher grades the likelihood to move into the duplication, idiosyncratic, or beginner analogical class at posttest decreases, and the likelihood to move into the advanced analogical class at posttest increases. In contrast, the higher the intensity of the training children received between pretest and posttest, the more likely they were to move into the duplication or idiosyncratic class at posttest, and the less likely they were to change into the intermediate analogical or advanced analogical class at posttest.

5. Discussion

The main aim of this paper was to provide a more detailed account of children's figural analogical reasoning development – more specifically we aimed to shed some light on which phases occur and which children transition when and how through these phases. We approached this by categorizing kindergarten to fourth grade children's (erroneous) solutions from a pretest-training-posttest study and analyzed the data with latent transition analyses. We now discuss three main findings: (1) analyses revealed five latent phases of analogical reasoning; (2) children generally progressed to more advanced reasoning phases from pretest to posttest; (3) age-group and how much feedback the child received during training were the clearest indicators of children's learning paths and rates of change, whereas working memory was only

Table 6

Regression coefficients of covariate effects on pretest phase occurrence and on transition probabilities, in beta (SE). Significant effects are in bold.

	Phase				
	1. Duplication	2. Idiosyncratic	3. Beginner	4. Intermediate	5. Advanced
Effects on pretest phase occurrence					
Grade	-0.96 (0.36)	-0.72 (0.37)	0.74 (0.35)	1.51 (0.35)	-0.58 (1.23)
WM	-0.97 (0.43)	-1.86 (0.45)	-0.73 (0.43)	0.35 (0.42)	3.20 (1.53)
Effects on probabilities to transition to each phase					
Grade	-4.76 (0.78)	-3.22 (0.66)	-2.34 (0.59)	-0.08 (0.63)	10.41 (2.02)
Training intensity	1.97 (0.40)	2.87 (0.45)	0.28 (0.30)	-2.13 (0.44)	-2.99 (0.71)

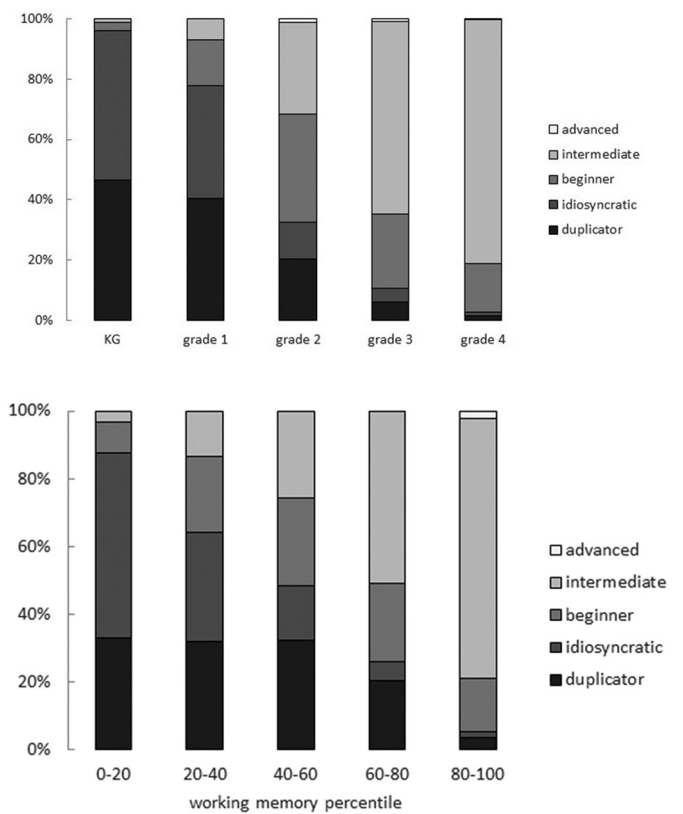


Fig. 7. Percentage of children in each class for each grade (upper panel) and working memory percentile range (lower panel) during pretest.

related to the children's reasoning phase at pretest, but not to their rate and path of change from pretest to posttest.

5.1. Latent phases in figural analogical reasoning

We hypothesized that there would be at least three main phases of figural analogical reasoning: a non-analogical phase characterized by mostly duplicate solutions; a transition phase where a mix solution types occur; and an analogical reasoning phase where duplicates are no longer given and analogies are consistently solved (partially) correctly. Instead, we identified five latent phases. We found two non-analogical reasoning phases, duplication and idiosyncratic. "Duplication" stood out as qualitatively different from the other phases as children in this phase solved almost all analogies by duplicating one of the matrix quadrants – irrespective of item difficulty; this was considered the least advanced phase and comprised 24% of children on the pretest and only 6% of children at posttest. The children's solutions in the "idiosyncratic" phase were also variable, but in this case partial solutions were common on easy items and other, idiosyncratic, solutions were generally provided on average and difficult items. In both cases, hardly any of the analogies were solved correctly. In contrast, the remaining three

phases showed few qualitative differences, instead they mostly differed with regard to how many items were solved (partially) correctly. In the “beginner analogical” reasoning phase, easy items were generally solved correctly, whereas average and difficult items were generally (partially) correct. Duplicate and other solutions also occurred in this phase, so there was some variability in solution type; therefore, this phase could perhaps be considered a transition phase where children transition from associative to relational reasoning strategies (e.g., [Hosenfeld et al., 1997](#); [Siegler & Svetina, 2002](#)). The two most advanced phases ‘intermediate analogical reasoning’ and ‘advanced analogical reasoning’ could be characterized by mostly partially correct or correct solutions on nearly all matrix analogies; in both cases, the children were clearly capable of solving the matrix analogies, but sometimes – especially on more difficult items in the intermediate analogical phase – they made mistakes (e.g., used the wrong sized animal) when solving them. None of the children were ‘advanced analogical reasoners’ at pretest and 36% solved most analogies correctly at posttest.

Previous work on children's matrix analogy solving often points to only two or three phases of reasoning. For example, [Siegler and Svetina \(2002\)](#) found that children generally progressed from a duplication phase to a phase with partially correct solutions in the trials leading up to the discovery of the how to correctly solve matrix analogies. [Hosenfeld et al. \(1997\)](#) found evidence of a similar transition from non-analogical to analogical reasoning. Our results reflect similar phases – non-analogical (duplication or idiosyncratic), transitioning to analogical (beginners), and analogical (intermediate and advanced); but, we have made further distinctions in the initial and later phases. Furthermore, our findings support [Siegler and Svetina's \(2002\)](#) conclusion that the transition from non-analogical to analogical reasoning is rather abrupt – in the “beginner” phase, the transition phase we identified, we see that children are already quite capable of solving analogies (partially) correct. Also, by using constructed-response items, rather than multiple-choice, we were able to demonstrate that children actually create duplicates, partially correct and various other solutions even if these are not offered. Furthermore, our results also show that the most variability in strategy-use occurs when transitioning (“beginner” phase), but not while duplicating or once figural analogy solving has been mastered, rather than throughout the acquisition of analogical reasoning (e.g., [Resing et al., 2017](#); [Siegler & Svetina, 2002](#); [Tunteler et al., 2008](#)).

5.2. Learning paths in figural analogical reasoning

When we look at the children's learning paths, we found, in line with our hypothesis, that children sequentially progressed towards more advanced analogical reasoning phases from pretest to posttest. But, there was still much variability in both the children's learning *rates* and *paths*. For example, some children stayed in one of the non-analogical phases, whereas others moved from a non-analogical phase at pretest to one of the most advanced phases, intermediate analogical reasoning, at posttest. Also, the only time children transitioned “down” a phase was from the idiosyncratic to the duplicating phase. Whenever children started in one of the analogical reasoning phases (beginner, intermediate, or advanced analogical) they either stayed in that phase or moved up, but not down. This implies that duplication and idiosyncratic reasoning phases can be seen as “entry” phases rather than phases children progress through when learning analogy solving. Relatedly, a qualitative change appears to take place in how children solve figural analogies when they progress from duplication or idiosyncratic reasoning to other phases – the shift from associative to relational strategies often noted in the literature ([Bulloch & Opfer, 2009](#); [Gentner, 1988](#)). Our results imply that some children make this shift in response to training, whereas others do not, which we discuss in further detail in the next paragraph. After this shift in reasoning occurs it appears that learning to solve figural analogies becomes more and more

refined, more of a quantitative shift, with increasingly more partial solutions being replaced by correct ones. In sum, our results show that analogical reasoning development cannot be captured by a one-size-fits-all pattern.

5.3. How age, working memory and training influence figural analogy solving development

We expected each of the three well-known accounts of analogical reasoning development, those of maturational, processing capacity and relational knowledge, to contribute to our understanding of which phases and learning paths children follow when learning to solve figural analogies. First, based on the maturational and knowledge accounts we expected children's analogical reasoning to improve with age. Indeed, the youngest children (Kindergarten and first graders) generally started off in the “duplication” phase. The “beginner analogical” or “intermediate analogical” reasoning phases at pretest contained mostly third and fourth graders. Second graders could be found in each of the phases at pretest. These age-groups correspond to those in the literature where 4–6 year-olds often require simplified tasks and additional instruction to solve analogies using relational rather than associative reasoning, whereas older children generally use relational information and are increasingly capable of correctly solving matrix analogies ([Chen et al., 2016](#); [Goswami & Brown, 1990a, 1990b](#); [Rattermann & Gentner, 1998](#)). Our cross-sectional observations mirror [Siegler & Svetina's](#) finding that age-related changes in analogical reasoning are similar to microgenetic changes within an age-group ([Siegler & Svetina, 2002](#)).

Children's age is a strong predictor of their ability to reason by analogy and our results are no exception. We also found that age was related to children's progression in figural analogy solving from pretest to posttest; more specifically, older children had “faster” learning trajectories and generally only the oldest age-groups, third and fourth graders, reached the most advanced phase of analogical reasoning. Previous microgenetic, dynamic testing, and training studies of analogical reasoning often included children within a small age range and/or did not include items that challenged older children ([Alexander et al., 1989](#); [Resing et al., 2017](#); [Siegler & Svetina, 2002](#); [Stevenson, Hickendorff, et al., 2013](#); [Tunteler et al., 2008](#)); we tested a relatively large age range and included items of varying difficulty. Consequently, our findings extend those of existing studies that examine the role of age in children's analogical reasoning development.

Second, based on the executive functions account, we expected the children's initial phase and path to be related to their performance on a working memory task. More specifically, children in more advanced phases and “faster” trajectories were expected to have higher working memory scores. Nearly 60% of the children in the “idiosyncratic” reasoning phase at pretest fell into the lowest working memory group, whereas approximately 80% of the children with above average working memory scores were “intermediate” analogical reasoners at pretest (the most advanced phase that occurred at that point). So, our findings show a positive relationship between working memory and analogical reasoning phase. This is in line with the literature showing that working memory plays a central role in cognitive and computational models of analogical reasoning ([Gentner & Forbus, 2011](#); [Halford, Wilson, & McDonald, 1995](#)) and is strongly related to children's initial analogy solving skills even after accounting for age ([Resing et al., 2017](#); [Stevenson, 2017](#); [Stevenson, Hickendorff, et al., 2013](#)). This also fits nicely with executive processing accounts of children's analogy solving development (e.g., [Richland et al., 2006](#); [Thibaut & French, 2016](#)). However, the influence of working memory on children's analogy solving progression in microgenetic and dynamic testing studies have produced contradictory results and demonstrated – at the least – that it is not the only mechanism driving analogical reasoning development ([Resing et al., 2017](#); [Ropovik et al., 2016](#); [Stevenson, 2017](#)). In contrast to expectations, working memory was not related to the children's learning paths, i.e., transitions from one phase to another,

in the current study. It could be that training somehow lowered the cognitive load of analogy solving, thereby neutralizing the effect of limited working memory, and leading to better analogy solving, which we will discuss next. However, in the future it would be important to examine this hypothesis by contrasting the effect of working memory on analogical reasoning development with and without training.

The third main mechanism that appears to steer analogical reasoning development is improved relational knowledge. When children learn to use abstract labels to refer to the relations between elements in an analogy they become more capable of solving analogies (Gentner, 2010). We compared two types of training, intelligent tutoring feedback (i.e., graduated prompts) and multiple-try feedback; one provided relational labels and the other did not. Surprisingly, the children's progression from pretest to posttest did not differ much between the two training groups. Only the amount of feedback they received affected their transition path; children who needed more feedback to solve the figural analogies were less likely to progress to more advanced reasoning phases, regardless of training type. It is important to note that the children in the multiple-try condition required more feedback in general than those in the graduated prompts condition. In addition, training intensity was related to both working memory and pretest performance so including this variable may be masking other effects. The results of this one study cannot outweigh the extant evidence of the relational knowledge account (Gentner, 2010; Gentner et al., 1995). But, it does add that relational knowledge may not need to be explicitly taught – perhaps multiple-try feedback stimulates children to change their approach, with which they independently transition to more advanced phases of analogical reasoning (e.g., Shute, 2008; Stevenson, 2017).

5.4. Limitations and future directions

A few limitations deserve mention when drawing conclusions about children's phases and learning paths in analogical reasoning based on these results. First, research from the executive functions “causal” account of the development of analogical reasoning suggest that young children often fail on analogical reasoning tasks because of their inability to inhibit associative responses (such as duplications or idiosyncratic solutions) (Richland et al., 2006; Thibaut & French, 2016). We aimed to minimize the influence of inhibition control by using constructed-response items where duplication and other solutions do not need to be inhibited. However, inhibition control likely still plays a role. A limitation of this study is that we did not directly measure inhibition control – this could perhaps explain which children remained duplicators or idiosyncratic reasoners and which ones progressed to (beginner, intermediate or advanced) analogical reasoning.

Second, there were only two measurement moments, one before and one after the training sessions. We could not include the training sessions in the analyses because the items were different from those used during testing. We recommend future research to use more testing sessions, each with the same or isomorphic items, to obtain a more complete picture of which children transition when to other phases in figural analogy solving.

Third, and finally, our study's sample size was somewhat limited. There were fifteen items at both pretest and posttest and each solution could be categorized into one of four categories. 388 participants is a small sample size for latent transition analyses. We did not have trouble estimating the different model parameters; however some of the hypothesized effects, such as the importance of working memory on transitions in the more advanced phases (Morrison et al., 2011), may not have been found due to lack of power. In the future, we advise obtaining a larger sample size to avoid this. Furthermore, we used two-hundred random starts when computing the final model in order to prevent local minima and recommend this in future work with latent transition analysis.

5.5. Conclusion

Our findings show that children's acquisition of figural analogical reasoning can be characterized by five phases showing great variability between and within children in the types of solutions they provided. Furthermore, although children generally progressed to more advanced phases from pretest to posttest, there was great variability in the rate and path of this progression through these phases. A critical strategy shift seems to take place where children learn to process and integrate multiple dimensions in the analogy rather than just copying one of the boxes from the matrix or using an idiosyncratic approach; this echoes Gentner's relational shift in the development of analogy solving (Gentner, 1988) and replicates the qualitative shift from non-analogical to analogical solutions in children's figural analogy solving (Hosenfeld et al., 1997; Siegler & Svetina, 2002). However, this shift does not seem to require maturation of working memory as previously hypothesized – in most children the shift can also be induced with training, both multiple-try and tutoring feedback helped. Latent transition analysis proved to be a very powerful tool to detect these patterns: a limited number of phases and transitions described this complex longitudinal dataset with its numerous categorical response variables. Our findings provide a rich addition to the literature, providing a more detailed depiction of the variability in how figural analogical reasoning develops.

References

- Alloway, T. P. (2007). *Automated working memory assessment*. London: Pearson Assessment.
- Alloway, T. P., Gathercole, S. E., & Pickering, S. J. (2006). Verbal and visuospatial short-term and working memory in children: Are they separable? *Child Development*, 77(6), 1698–1716. <http://dx.doi.org/10.1111/j.1467-8624.2006.00968.x>.
- Alexander, P. A., Jablansky, S., Singer, L. M., & Dumas, D. (2016). Relational reasoning: What we know and why it matters. *Policy insights from the behavioral and brain sciences* <http://dx.doi.org/10.1177/2372732215622029>.
- Alexander, P. A., Willson, V. L., White, C. S., Fuqua, J. D., Clark, G. D., Wilson, A. F., & Kulikowich, J. M. (1989). Development of analogical reasoning in 4- and 5-year-old children. *Cognitive Development*, 4(1), 65–88. [http://dx.doi.org/10.1016/0885-2014\(89\)90005-1](http://dx.doi.org/10.1016/0885-2014(89)90005-1).
- Bouwmeester, S., Sijtsma, K., & Vermunt, J. K. (2004). Latent class regression analysis for describing cognitive developmental phenomena: An application to transitive reasoning. *European Journal of Developmental Psychology*, 1(1), 67–86. <http://dx.doi.org/10.1080/17405620344000031>.
- Bullock, M. J., & Opfer, J. E. (2009). What makes relational reasoning smart? Revisiting the perceptual-to-relational shift in the development of generalization. *Developmental Science*, 12(1), 114–122. <http://dx.doi.org/10.1111/j.1467-7687.2008.00738.x>.
- 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, New York, USA: Guilford Press.
- Chen, Z., Honomichl, R., Kennedy, D., Tan, E., Chen, Z., Honomichl, R., ... Tan, E. (2016). Aiming to complete the matrix: Eye-movement analysis of processing strategies in children's relational thinking aiming to complete the matrix: Eye-movement analysis of processing strategies in children's relational thinking. *Developmental Psychology*, 52(6), 867–878.
- Cheshire, A., Muldoon, K. P., Francis, B., Lewis, C. N., & Ball, L. J. (2007). Modelling change: New opportunities in the analysis of microgenetic data. *Infant and Child Development*, 16(1), 119–134. <http://dx.doi.org/10.1002/icd.498>.
- Collins, L. M., & Lanza, S. T. (2010). *Latent class and latent transition analysis: With applications in the social sciences*. Hoboken, NJ, USA: John Wiley & Sons, Inc. <http://dx.doi.org/10.1002/9780470567333>.
- Elliott, J. G., Grigorenko, E. L., & Resing, W. C. M. (2010). Dynamic assessment. *International Encyclopaedia of Education* (pp. 220–225). (3rd ed.). Elsevier.
- Engel de Abreu, P. M. J., Conway, A. R. A., & Gathercole, S. E. (2010). Working memory and fluid intelligence in young children. *Intelligence*, 38(6), 552–561. <http://dx.doi.org/10.1016/j.intell.2010.07.003>.
- Gentner, D. (1988). Metaphor as structure mapping: The relational shift. *Child Development*, 59(1), 47. <http://dx.doi.org/10.2307/1130388>.
- Gentner, D. (2010). Bootstrapping the mind: Analogical processes and symbol systems. *Cognitive Science*, 34, 752–775.
- Gentner, D., & Forbus, K. D. (2011). Computational models of analogy. *Wiley Interdisciplinary Reviews: Cognitive Science*, 2(3), 266–276. <http://dx.doi.org/10.1002/wcs.105>.
- Gentner, D., Rattermann, M. J., Markman, A., & Kotovsky, L. (1995). Two forces in the development of relational similarity. In G. Halford, & T. Simon (Eds.), *Developing cognitive competence: New approaches to process modeling* (pp. 263–313). Earlbaum: Hillsdale.
- Goswami, U. (1989). Relational complexity and the development of analogical reasoning.

- Cognitive Development*, 4(3), 251–268. [http://dx.doi.org/10.1016/0885-2014\(89\)90008-7](http://dx.doi.org/10.1016/0885-2014(89)90008-7).
- Goswami, U. (1991). Analogical reasoning: What develops? A review review of research and theory. *Child Development*, 62, 1–22.
- Goswami, U., & Brown, A. L. (1990a). Higher-order structure and relational reasoning: Contrasting analogical and thematic relations. *Cognition*, 36(3), 207–226. [http://dx.doi.org/10.1016/0010-0277\(90\)90057-Q](http://dx.doi.org/10.1016/0010-0277(90)90057-Q).
- Goswami, U., & Brown, A. L. (1990b). Melting chocolate and melting snowmen: Analogical reasoning and causal relations. *Cognition*, 35(1), 69–95. [http://dx.doi.org/10.1016/0010-0277\(90\)90037-K](http://dx.doi.org/10.1016/0010-0277(90)90037-K).
- Grossnickle, E. M., Dumas, D., Alexander, P. A., & Baggetta, P. (2016). Individual differences in the process of relational reasoning. *Learning and Instruction*, 42, 141–159.
- Halford, G. S., Wilson, W. H., & McDonald, M. (1995, July). Complexity of structure mapping in human analogical reasoning: A PDP model. *Proceedings of the seventeenth annual conference of the cognitive science society*.
- Halford, G. S., Wilson, W. H., & Phillips, S. (1998). Processing capacity defined by relational complexity: Implications for comparative, developmental, and cognitive psychology. *Behavioral and Brain Sciences*, 21(6), 803–831.
- Hickendorff, M., Edelsbrunner, P. A., McMullen, J., Schneider, M., & Trezise, K. (2017). Informative tools for characterizing individual differences in learning: Latent class, latent profile, and latent transition analysis. *Learning and Individual Differences*. <http://dx.doi.org/10.1016/j.lindif.2017.11.001>.
- Hornung, C., Brunner, M., Reuter, R. A. P., & Martin, R. (2011). Children's working memory: Its structure and relationship to fluid intelligence. *Intelligence*, 39(4), 210–221. <http://dx.doi.org/10.1016/j.intell.2011.03.002>.
- Hosenfeld, B., van der Maas, H. L. J., & van den Boom, D. C. (1997). Indicators of discontinuous change in the development of analogical reasoning. *International Journal of Behavioral Development*, 20, 367–395. <http://dx.doi.org/10.1006/jecp.1996.2351>.
- Kail, R. V. (2007). Longitudinal evidence that increases in processing speed and working memory enhance children's reasoning. *Psychological Science*, 18(4), 312–313. <http://dx.doi.org/10.1111/j.1467-9280.2007.01895.x>.
- Melby-Lervåg, M., & Hulme, C. (2013). Is working memory training effective? A meta-analytic review. *Developmental Psychology*, 49(2), 270–291.
- Morrison, R. G., Dumas, L. A., & Richland, L. E. (2011). A computational account of children's analogical reasoning: Balancing inhibitory control in working memory and relational representation. *Developmental Science*, 14(3), 516–529.
- Mulholland, T. M., Pellegrino, J. W., & Glaser, R. (1980). Components of geometric analogy solution. *Cognitive Psychology*, 12(2), 252–284. Retrieved from: <http://www.ncbi.nlm.nih.gov/pubmed/7371379>.
- Nylund, K. L., Asparouhov, T., & Muthén, B. O. (2008). Deciding on the number of classes in latent class analysis and growth mixture modeling: A Monte Carlo simulation study: Erratum. *Structural Equation Modeling*, 15(1), 182. <http://dx.doi.org/10.1080/10705510701793320>.
- Rattermann, M. J., & Gentner, D. (1998). More evidence for a relational shift in the development of analogy: Children's performance on a causal-mapping task. *Cognitive Development*, 13(4), 453–478. [http://dx.doi.org/10.1016/S0885-2014\(98\)90003-X](http://dx.doi.org/10.1016/S0885-2014(98)90003-X).
- Resing, W. C. M., Bakker, M., Pronk, C. M. E., & Elliott, J. G. (2017). Progression paths in children's problem solving: The influence of dynamic testing, initial variability, and working memory. *Journal of Experimental Child Psychology*, 153, 83–109. <http://dx.doi.org/10.1016/j.jecp.2016.09.004>.
- 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. *The British Journal of Educational Psychology*, 81(Pt 4), 579–605. <http://dx.doi.org/10.1348/2044-8279.002006>.
- Richland, L. E., & Burchinal, M. R. (2013). Early executive function predicts reasoning development. *Psychological Science*, 24(1), 87–92.
- Richland, L. E., Morrison, R. G., & Holyoak, K. J. (2006). Children's development of analogical reasoning: Insights from scene analogy problems. *Journal of Experimental Child Psychology*, 94(3), 249–273. <http://dx.doi.org/10.1016/j.jecp.2006.02.002>.
- Ropovik, I., Bobáková, M., Ferjenčík, J., Filičková, M., Kovalčíková, I., & Slavkovská, M. (2016). Explaining the ability to learn analogies: the role of executive functions and fluid intelligence. *Stud. Psychol.* 58(4), 322–335. <http://dx.doi.org/10.21909/sp.2016.04.726>.
- Shute, V. J. (2008). Focus on formative feedback. *Review of Educational Research*, 78(1), 153–189. <http://dx.doi.org/10.3102/0034654307313795>.
- Siegler, R. S., & Svetina, M. (2002). A microgenetic/cross-sectional study of matrix completion: Comparing short-term and long-term change. *Child Development*, 73(3), 793–809. <http://dx.doi.org/10.1111/1467-8624.00439>.
- Sternberg, R. J., & Rifkin, B. (1979). The development of analogical reasoning processes. *Journal of Experimental Child Psychology*, 27(2), 195–232. [http://dx.doi.org/10.1016/0022-0965\(79\)90044-4](http://dx.doi.org/10.1016/0022-0965(79)90044-4).
- Stevenson, C. E. (2017). Role of working memory and strategy-use in feedback effects on children's progression in analogy solving: An explanatory item response theory account. *International Journal of Artificial Intelligence in Education*, 27(3), 393–418.
- 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(3), 159–169. <http://dx.doi.org/10.1016/j.cedpsych.2013.02.001>.
- Stevenson, C. E., Heiser, W. J., & Resing, W. C. M. (2016). Dynamic testing of analogical reasoning in 5- to 6-year-olds: Multiple-choice versus constructed-response training items. *Journal of Psychoeducational Assessment*, 34(6), 550–565. <http://dx.doi.org/10.1177/0734282915622912>.
- Stevenson, C. E., Hickendorff, M., Resing, W. C. M., Heiser, W. J., & de Boeck, P. A. L. (2013). Explanatory item response modeling of children's change on a dynamic test of analogical reasoning. *Intelligence*, 41(3), 157–168.
- Thibaut, J.-P., & French, R. M. (2016). Analogical reasoning, control and executive functions: A developmental investigation with eye-tracking. *Cognitive Development*, 38, 10–26. <http://dx.doi.org/10.1016/j.cogdev.2015.12.002>.
- Tunteler, E., Pronk, C. M. E., & Resing, W. C. M. (2008). Inter- and intra-individual variability in the process of change in the use of analogical strategies to solve geometric tasks in children: A microgenetic analysis. *Learning and Individual Differences*, 18(1), 44–60. <http://dx.doi.org/10.1016/j.lindif.2007.07.007>.
- Tunteler, E., & Resing, W. C. (2002). Spontaneous analogical transfer in 4-year-olds: A microgenetic study. *Journal of Experimental Child Psychology*, 83(3), 149–166. [http://dx.doi.org/10.1016/S0022-0965\(02\)00125-X](http://dx.doi.org/10.1016/S0022-0965(02)00125-X).
- Tzuril, D., & George, T. (2009). Improvement of analogical reasoning and academic achievement by the analogical reasoning programme (ARP). *Educational and Child Psychology*, 26(3), 71.
- Tzuril, D., & Klein, P. S. (1985). The assessment of analogical thinking modifiability among regular, special education, disadvantaged, and mentally retarded children. *Journal of Abnormal Child Psychology*, 13(4), 539–552.
- Vermunt, J. K., & Magidson, J. (2013a). *Latent GOLD 5.0. [Computer Software]*. Belmont Massachusetts: Statistical Innovations Inc.
- Vermunt, J. K., & Magidson, J. (2013b). *Technical guide for latent GOLD 5.0: Basic, advanced, and syntax*. Belmont Massachusetts: Statistical Innovations Inc.