EXPLAINING THE ABILITY TO LEARN ANALOGIES: THE ROLE OF EXECUTIVE FUNCTIONS AND FLUID INTELLIGENCE

Ivan ROPOVIK[1], Monika BOBÁKOVÁ[1], Ján FERJENČÍK[2], Marta FILIČKOVÁ[1], Iveta KOVALČIKOVÁ[1], Miriam SLAVKOVSKÁ[2]

[1]Department of Preschool and Elementary Education and Psychology
Faculty of Education, University of Prešov
17. Novembra 15, 080 01 Prešov, Slovak Republic
E-mail: ivan.ropovik@unipo.sk

[2] Department of Psychology, University of Pavol Jozef Šafárik

Abstract: It is well established that analogical reasoning can be explained by the efficiency of working memory (WM) but it remains unclear what processes are involved when the child learns to reason analogically. The present study examined the relationship of executive functions (EF) and fluid intelligence (gF) and the ability to learn analogies in a sample of 210 10-year-old children. First, with regard to the structure of EF, a four-factor model fitted the data well, however, shifting and fluency were indistinguishable from attentional control. At the same time, attentional control fully accounted for the interrelationships between other EF. Second, only WM proved to have a direct effect on the ability to learn and on gF, while mediating the effect of attentional control. Third, despite a decent explanatory power of WM, it did not explain the relationship between the ability to learn and gF, indicating the presence of another factor distinct from WM.

Key words: ability to learn, analogical reasoning, fluid intelligence, executive functions, working memory

Explaining the Ability to Learn Analogies: The Role of Executive Functions and Fluid Intelligence

Much of learning requires relating different concepts and transferring knowledge from a well understood domain to one that is unfamiliar (Bransford, Franks, Yve, & Sherwood, 1989; Goswami, 1992). Analogical reasoning is thus one of the most impor-

tant abilities involved in making inferences about new phenomena, learning how to solve novel situations, and extracting relevant information from an experience on the basis of relational similarity (Chen, Sanchez, & Campbell, 1997; Richland, Morrison, & Holyoak, 2006). Past research suggests the existence of three factors underlying the age-related differences in analogical reasoning (see Richland et al., 2006).

First, according to Goswami (1992), children as young as 3 years are able to derive correct analogies, provided they possess the relevant pre-existing domain knowledge. In contrast to the Piagetian notion of succes-

sive developmental stages, it may only be the lack of relevant conceptual knowledge that poses the primary constraint for analogical reasoning during early years. Second, some studies propose the existence of a "relational shift" from featural to relational similarities. According to this hypothesis, attribute matching precedes relation encoding while the developmental preference for the latter is context dependent and is driven by the knowledge of the given domain (Rattermann & Gentner, 1998). Third, as repeatedly shown, performance in analogical reasoning tasks can be explained by the efficiency of working memory (WM). Analogical reasoning requires extracting, maintaining and manipulating multiple relations simultaneously. Because these relations need to be processed at the same time, there is an inherent need for a system that builds relational representations through temporary bindings between component representations (Oberauer, Süß, Wilhelm, & Wittmann, 2008), i.e., the WM. Indeed, WM seems to be strongly related to the ability of analogical reasoning (Cho, Holyoak, & Cannon, 2007) and to inductive reasoning in general (Kyllonen & Christal, 1990). Moreover, the processes involved in mapping multiple relations have been shown to be mediated by the same areas of prefrontal cortex as WM processes (Kroger, 2002).

Present Study

Although there is quite a solid body of evidence regarding the processes underlying analogical reasoning as a static ability in discrete developmental stages, we still do not know much about the dynamic aspect of analogical reasoning, i.e., what processes are involved when the child actually learns to reason analogically. The question here is, what processes drive the individual differences in the development of analogical reasoning on a micro level within a specific learning situation?

Analogical reasoning is itself a complex cognitive process and the additional demands associated with higher-order learning (internalizing the principles and acquiring novel response routines) induce the need for some regulatory processes that usually fall under the umbrella term executive functions (EF). EF refer to a family of top-down mental functions that control and organize mental processes and include functions like response inhibition and interference control, (set-)shifting, or WM (Diamond, 2013). Although, as mentioned above, the effect of WM on analogical reasoning has already been studied, it is not clear what role it takes when embedded within an explanatory structure of other relevant EF and fluid intelligence (gF). With regard to the structure of executive functioning and its dynamics during development, there is now a large body of evidence. Until approximately 9 years of age, diverse aspects of executive functioning have repeatedly been shown to follow a single dimension (Brydges, Reid, Fox, & Anderson, 2012; Wiebe, Espy, & Charak, 2008). A bit later, at the onset of adolescence, the former unitary executive functioning is consistently found to be manifested in diverse mental functions like inhibition, shifting, or WM (Lehto, Juujärvi, Kooistra, & Pulkkinen, 2003; Miyake et al., 2000) which, in turn, give rise to higher-order functions like problem-solving or planning (Klenberg, Korkman, & Lahti-Nuuttila, 2001). However, there is still a lack of convincing evidence about the exact structure of EF between these two developmental stages, nor is it clear

which aspect of executive functioning dissociates first, possibly driving the differentiation process further on.

The objective of the present study was to examine the structure of relationships between 1) two, supposedly diverse aspects of fluid mental ability, i.e., the ability to learn in the domain of analogical reasoning (dynamic aspect, denoted as the "ability to learn analogies") and fluid intelligence (static aspect, "gF"), and 2) four postulated EF, namely "attentional control", "fluency", "shifting" and "WM"). The here proposed theory, formally defined by means of a structural equation model, laid down a set of hypotheses. First, it was tested, whether the executive functioning can be represented already at the age of 10 years by a structure of four diverse functions, as mentioned above. Second, it was hypothesized that the attentional control domain (involving interference control and response inhibition aspect) fully accounts for any uni- or bi-directional relationships between all other EF defined within the model. In line with past research the functioning of top-down regulatory mechanisms (i.e., the EF) was expected to still strongly rely on shared pool of attentional resources (Cowan, Morey, Chen, & Bunting, 2007; Engle, Tuholski, Laughlin, & Conway, 1999), due to yet incomplete developmental shift to a system of related, but rather diverse functions. Is the attentional control still such a strong factor that it can explain the relationships between the other EF? Third, it was predicted that WM mediates the effect of attentional control on gF and the ability to learn analogies. At the same time, it was expected that neither shifting, nor fluency exerts a direct effect on the mentioned target variables above and beyond the effect of WM (see Ropovik, 2014). Fourth, it was

tested whether variation in WM fully explains the relationship between gF and the ability to learn analogies.

Method

Participants and Procedure

The participants were 210 Caucasian (Slovak) children attending the last (4th) grade of elementary school, 124 girls and 86 boys, with a mean age of 9 years and 9 months ($SD = 6.5$ months; $IQR = 112 - 122$ months). The selection of subjects for the sample employed the cluster sampling technique, with the entire classes of elementary schools representing the clusters. Based on 2011 census data, 12 classes (clusters) of elementary schools were selected, proportionally stratified by the size of residence (into three levels). The mean size of a cluster was 17.5 children. Child's participation in the study was conditioned on obtaining informed consent from parents.

Every child was tested individually by trained psychologists on the measures described below. Testing took place before noon in a quiet room on three occasions, lasting approximately 180 minutes total.

Measures

The employed measures were selected from three test batteries, the Delis-Kaplan Executive Function System (Delis, Kaplan, & Kramer, 2001), Woodcock-Johnson International Editions (Ruef, Furman, & Muñoz-Sandoval, 2003), and AnimaLogica (Stevenson, Hickendorff, Resing, Heiser, & De Boeck, 2013). Two indicators per every defined latent variable were used in order to alleviate the task impurity problem. This is

essential especially with EF since it is not possible to measure them in isolation of other non-executive cognitive functions (Anderson, 2002). This frequently leads to low reliability estimates due to the overrepresentation of construct-irrelevant variance (Rabbitt, 1997) and the consequent inability to test substantive hypotheses based on attenuated correlations of these measures within the Classical test theory. With regard to the input and response modalities, conceptual complexity and non-executive characteristics, the tasks were chosen to be as divergent as possible to make sure that the effect of the indicated latent variable is the most likely explanation of the shared variance.

Visual Matching (W-J IE). A cancellation task measuring mental speed, where the task was to identify and mark matching numbers in a series of six numbers. The measure was used as one of the indicators of attentional control.

Color-Word Interference Test (D-KEFS). This rendition of the Stroop task required inhibition of overlearned response, i.e., the conflicting response to stimuli with incongruent features (meaning and ink color), the Inhibition condition (Subtest 3) score was used as an indicator of attentional control.

Verbal fluency test (D-KEFS). Two of the Verbal fluency subtests were employed, the Letter fluency and Switching conditions. Here, the subject was required to produce as many words as possible – within a 60s time limit and under restricted search conditions (words beginning only with a certain letter or belonging to two defined categories to be switched between). The total number of given words in Letter fluency served as a measure of fluency, while the total number of correct switches in the Switching condition was intended to provide a measure of the shifting factor. The rank-order correlation between these two conditions was $r_s = .24, p < .001$.

Design fluency test (D-KEFS). In this nonverbal fluency measure, the task was to generate as many novel abstract designs in 60s as possible by connecting dots. The total number of correct designs across the three subtests provided the dependent measure of nonverbal fluency, further used as an indicator of fluency.

Trail making test (D-KEFS). The tasks relevant for this study included linking numbers with letters in alternating order (Subtest 4). The time taken to complete the task served as an indicator of shifting.

Numbers Reversed (W-J IE). A measure of the verbal aspect of WM that required the subjects to repeat numbers in reverse order.

Tower test (D-KEFS). A complex task that requires the subject to inhibit prepotent responses, devise a solution plan, hold it in WM and monitor performance. The task was to move disks across three pegs according to rules to reach a given goal state with as few moves as possible. The total number of rule violation moves was used as the dependent measure of WM in visuo-spatial domain.

Spatial Relations (W-J IE). This test measures the visuo-spatial thinking aspect of gF. The task was to choose two or three shapes that make-up the target abstract shape.

Quantitative Reasoning (W-J IE). A measure that directly taps the gF, requiring the deduction of quantitative concepts and principles.

Verbal analogies (W-J IE). A measure of the ability to identify verbal relationships. The task was to complete three-word analogies in the form A:B::C:D. The raw test score was used to derive an unstandardized re-

sidual score indicating the ability to learn analogies.

AnimaLogica. A fully computerized dynamic test measuring the ability to learn in the domain of figural analogies. The measure employs a pretest-intervention-posttest design, in which pretest and posttest (20 items each) were designed as isomorphic measures with no help provided (Stevenson, Bergwerff, Heiser, & Resing, 2014). The intervention (teaching) phase followed the graduated-prompt procedure (Campione & Brown, 1987) that was based on a series of five hints (from metacognitive through cognitive to solution constructing prompts), progressively revealing the solution in each of the ten analogy items. Within a 2 x 2 matrix, the subject was required to place the missing animal figures in order to complete the analogy. There were between two to eight variations of the animal figures according to their number, size, color, orientation, and position. The following two indicators of the ability to learn analogies were used. To isolate the "ability to learn" component, unstandardized residual score was used, such that the variance in AnimaLogica raw posttest score, which accounted for by the ability of analogical reasoning (measured by Verbal Analogies test), was removed. The other dependent measure was the number of prompts that children needed in achieving successful independent performance within the learning phase, reflecting the revelation of the child's zone of proximal development (Vygotsky, 1987).

Data Analysis

In order to test the above defined set of hypotheses and see if the proposed theoretical structure fits the data, structural equation modeling (SEM) was used (model outlined in Table 3). The analysis of the covariance matrix was conducted using the maximum likelihood estimation method in AMOS 22.0 (Arbuckle, 2013). Given an expected adequate statistical power (computed in R) with regard to the specific model complexity and sample size, a significant χ^2 value ($p < .05$) was regarded a sufficient criterion for model rejection, irrespective of the approximate goodness-of-fit indices (see Hayduk, 2014; Ropovik, 2015). For a non-rejected model, the following approximate indices were further examined: the Comparative Fit Index (*CFI*), the Tucker-Lewis Index (*TLI*), the Root Mean Square Error of Approximation (*RMSEA*), the Standardized Root Mean Square Residual (*SRMR*) and the Bayesian Information Criterion (*BIC*). The usually suggested "rules of thumb" cut-off criteria indicating a well-fitting model were followed: *CFI* and *TLI* > .95, *RMSEA* < .06 and *SRMR* < .08 (Hu & Bentler, 1999). Due to the non-parametric (skewed and leptokurtic) character of some indicators (see Table 1), the Bollen-Stine bootstrap (Bollen & Stine, 1992) was used to estimate the standard errors of model parameters and to correct the model test (χ^2 test) significance (using 2000 samples). The estimates of model parameters were interpreted only in case of no evidence of global or local model misspecification.

Results

Data Screening

Prior to the data analysis, the data were screened for normality, missing or improbable values, and univariate outliers. Multiple imputation method was used to handle the missing data (0.2%). The variables were

checked for outlying values based on a matrix of z-scores. If the distribution contained more than 3 excessive values ($x > M \pm 2SD$), outlying cases were assigned a raw score that was one unit larger (or smaller) than the next most extreme score in the distribution of the offending variable (Tabachnick & Fidell, 2007). All variables were tested for gender, age, and clustering effects. There were no significant differences for gender. With regard to age (two levels split by median age), significant differences (in favor of the older group) were found for the following variables: Visual Matching, C-W Interference Test, Design Fluency, TMT Switching, Numbers Reversed, Tower Test, Spatial Relations, Quantitative Reasoning. The effect sizes (r) were, however, of rather small magnitude: .26, .19, .17, .15, .15, .15, .14, and .16, respectively. No significant intra-class correlations indicating a cluster effect (Bonferroni corrected) on any of the variables were found. Descriptive statistics for the raw scores can be seen in Table 1. No nonlinear transformations were performed. The matrix of zero-order correlations for the indicator variables is presented in Table 2.

Model Testing

The proposed set of hypotheses was tested within a structural equation model, as defined in Table 3. Fitting of the model with $df = 49$ to the sample covariance matrix converged to an admissible solution without any convergence problems. However, the model did not fit the data well enough, given the $\chi^2 = 76.1$ and the associated $p = .01$. Subsequent model diagnostics (residual covariances, modification indices, exploratory analyses of the measurement models) revealed two model misspecifications, leading

to the following theory-driven changes to the initial model. The first misspecification concerned the measurement model of attentional control, reflecting the inability of the superordinate latent variable to fully explain the covariance of its respective indicators measuring two aspects of attentional control, namely interference control (Visual Matching) and response inhibition (C-W Interference Test – Inhibition). This fact can be explained in two ways. Apart from the power aspect (the factor of attentional control), the performance on these tasks could be also excessively affected by mental speed. Alternatively, the two subsystems of attentional control (interference control and response inhibition) may already be regarded interrelated but separate constructs at the given age (Ropovik, 2014). The second misspecification concerned the structural model and spoke against the formulated hypothesis that WM fully accounts for the relationship between gF and the ability to learn analogies. Here, the data suggest the presence of another important factor at play that is distinct from WM.

The initial model was respecified in order to comply with the changes to the above discussed hypotheses, namely by 1) adding an error covariance between the indicators of attentional control and 2) modeling a covariance between the disturbance terms of gF and the ability to learn analogies (see Figure 1, indicated by a dashed line). Given the model test, the respecified model provided a good fit to the data ($\chi^2 = 61.5$; $df = 47$; $p = .08$). The values of the approximate fit indices were favorable as well with $CFI = .97$; $TLI = .96$; $RMSEA = .038$, 90% CI [.00, .06]; $SRMR = .048$; and $BIC = 227$. With regard to the local fit, the matrix of standardized residuals was inspected for significant

Table 1 *Descriptive Statistics*

	M	*SD*	Skewness	Kurtosis	*Min*	*Max*
Visual Matching	36.29	5.98	-.35	.02	18	52
C-W Interference Test	84.86	26.60	1.40	2.23	40	180
VF Letter Fluency	19.09	7.05	.57	.29	4	44
Design Fluency	19.93	5.12	.35	-.33	10	35
TMT Switching	117.00	45.42	.94	.44	31	240
VF Switching	9.19	3.07	-.28	.35	0	17
Numbers Reversed	13.60	3.33	.40	.26	5	24
Tower Test	16.12	3.39	.04	.51	6	26
Spatial Relations	20.31	3.77	.60	-.29	12	31
Quantitative Reasoning	26.09	3.73	-.50	.49	13	34
AnimaLogica Residual	.00	1.00	-1.38	4.86	-5.75	1.94
AnimaLogica Hints	3.74	3.42	1.98	6.73	-2.75	24.00

Note. VF = Verbal Fluency; TMT = Trail Making Test; $SE_{Skewness}$ = 0.17; $SE_{Kurtosis}$ = 0.33

Table 2 *Correlations*

	[1]	[2]	[3]	[4]	[5]	[6]	[7]	[8]	[9]	[10]	[11]
[1] Visual Matching	1.00										
[2] C-W Interference Test	-.50**	1.00									
[3] VF Letter Fluency	.24**	-.22**	1.00								
[4] Design Fluency	.23**	-.29**	.24**	1.00							
[5] TMT Switching	-.40**	.42**	-.28**	-.33**	1.00						
[6] VF Switching	.20**	-.32**	.23**	.19**	-.35**	1.00					
[7] Numbers Reversed	.30**	-.24**	.26**	.12	-.27**	.16*	1.00				
[8] Tower Test	.14*	-.19**	-.03	.07	-.28**	.18*	.15*	1.00			
[9] Spatial Relations	.17*	-.11	.17*	.18**	-.18**	.06	.36**	.05	1.00		
[10] Quantitative Reasoning	.23**	-.09	.27**	.22**	-.33**	.14*	.45**	.21**	.45**	1.00	
[11] AnimaLogica Residual	.19**	-.14*	.19**	.15*	-.22**	.08	.27**	.07	.25**	.35**	1.00
[12] AnimaLogica Hints	-.05	.13	-.15*	-.12	.18**	-.15*	-.19**	-.06	-.30**	-.38**	-.56**

Note. Pearson correlations; ** p < .01; * p < .05 (2-tailed); VF = Verbal Fluency; TMT = Trail Making Test

Table 3 *Specification of the initial structural equation model*

Measurement model	Structural model
LV1 Attentional control BY Visual Matching, C-W Interference Test [Inhibition].	*LV2, LV3, LV4* ON
LV2 Fluency BY Verbal Fluency Test [Letter Fluency], Design Fluency.	*LV1;*
LV3 Shifting BY Trail Making Test [Switching], Verbal Fluency [Switching].	*LV5, LV6* ON *LV4.*
LV4 WM BY Numbers Reversed, Tower Test.	
LV5 gF BY Spatial Relations, Quantitative Reasoning.	
LV6 Ability to learn analogies BY AnimaLogica Residual, AnimaLogica Hints.	

Note. (defined) BY; (regressed) ON; LV = latent variable
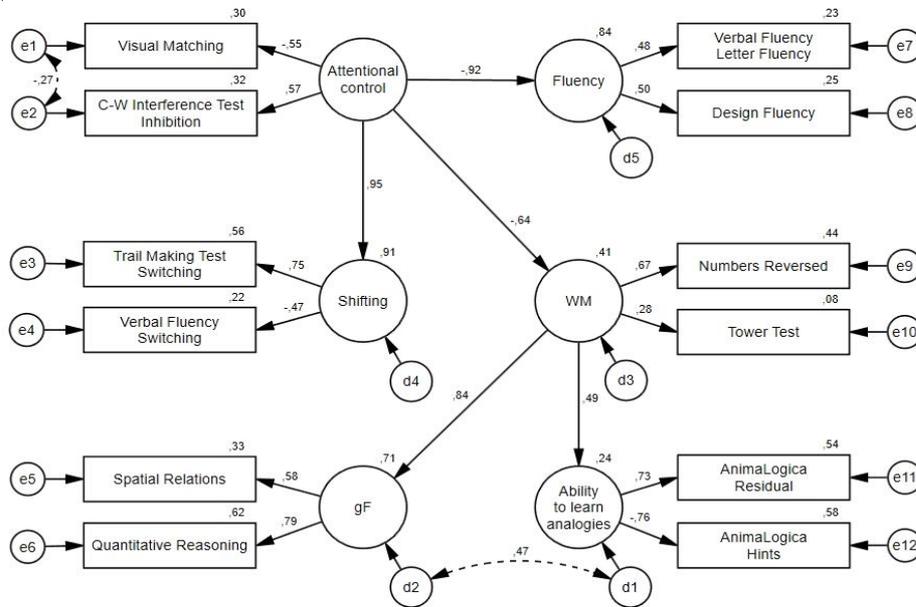
Figure 1 The estimated structural equation model. Negative values reflect inverse scaling of the given variable.

residual covariances. Out of the 78 matrix elements, two residuals crossed the threshold value of ± 2*SE* with *z*-values of 2.1 and 2.2 but, following a detailed review of the involved variables, both indications of local misfit were eventually deemed practically negligible. To estimate the power for the test of close-fit hypothesis, we followed the approach by MacCallum, Browne, and Sugawara (1996), which is based on the distribution of the *RMSEA*. The aim was to determine the likelihood of rejecting the conclusion that the model provides a close fit when it actually does not. For the hypothesized model with $df = 47$ and $N = 210$, there was an adequate power of .77 to uncover

beyond chance model-data discrepancies from near perfect model fit and reject an incorrect model if it were the case ($H0$: $\varepsilon \leq .05$, $\varepsilon_1 = .08$).

In order to reduce the possibility that a different factorial structure of EF explains the observed covariances in a more efficient way, an alternative model was tested. Based on the notion of Brydges et al. (2012), the model defined executive functioning as a unitary construct, where the performance in all of the EF indicators fell along a single dimension. Such a factorial structure embedded within the full structural model was, however, not supported by the data ($\chi^2 = 84.3$; $df = 50$; $p = .002$). At the given age, executive

functioning thus seems to be better represented by a set of highly related, but already diverse mental functions.

Model Evaluation

Backed up by adequate statistical power, the χ^2 test was not able to formally reject the hypothesis of correct model specification. Since the model is likely to reproduce the observed empirical relations well, it is justified to interpret the estimated parameters, which would not have been the case had the model test failed (Antonakis, Bendahan, Jacquart, & Lalive, 2010). The diagram of the estimated structural model can be seen in Figure 1. With respect to the measurement models, the indicators of the four latent EF had rather low factor loadings (ranging from .75 to .28). The average variance extracted (*AVE*) for the attentional control, fluency, shifting, and WM was at .31, .24, .39, and .26, respectively. Although expected, it has to be noted that such low values of extracted variance fall far behind the customary psychometric criteria (e.g., $AVE \geq .50$). As usual in the research of executive functioning, the construct identity of the formulated latent variables is consequently a bit on shaky ground. On the other hand, the indicators of gF and the ability to learn analogies did better in measuring the respective latent variables with *AVE* of .48 and .56 for gF and the ability to learn analogies, respectively.

Whereas the measurement models represent a fundamental psychometric level, the formulated hypotheses concerned primarily the latent variable level and the interrelationships within the structural model. As shown above, the adequate fit of the respecified model speaks in favor of the first hypothesis, defining four distinct EF. At the same

time, attentional control fully accounted for all the relationships between other executive functions, as predicted by the second hypothesis. However, a closer look at the regression path coefficients made it obvious that attentional control, fluency and shifting were highly collinear. Although such a model was not hypothesized a priori, merging attentional control, fluency and shifting into a single construct reflecting general mental efficiency provided a good fit to the data. With $df = 49$, model test yielded a χ^2 of 62.2 with the associated probability $p = .098$ and almost identical approximate fit indices ($CFI = .97$; $TLI = .96$; $RMSEA = .036$, 90% CI [.00, .06]; $SRMR = .048$; and $BIC = 217$). This exploratory model involving a two-factor EF structure (nested within the more complex model involving a four-factor EF structure) was a bit more parsimonious, and it provided an identical fit ($\chi^2_{Diff}(2) = 0.7$, $p = .71$). Regarding the relationship to WM, the regression path from the merged attentional control/mental efficiency to WM remained practically unchanged, at -.63. However, because of the fact that models relying heavily on a posteriori data-driven modification frequently capitalize on chance variation, making them usually irreproducible (MacCallum, Roznowski, & Necowitz, 1992), we stick to the model involving a four-factor structure of EF for interpretation.

With regard to the third hypothesis, the data provided evidence for the direct effect of WM on the ability to learn analogies (.49) as well as on gF (.84). However, other EF lacked a direct link to these target constructs. Here, all the effect of attentional control on the ability to learn analogies or gF was mediated by WM, with indirect effects of .54 and .31, respectively. After controlling for the effect of WM, none of the other executive func-

tions were able to explain a significant proportion of variance in gF and the ability to learn analogies.

Over and above the expected explanatory power of WM with regard to gF and the ability to learn analogies, the fourth hypothesis predicted that WM can account for all the shared variance between those target constructs. However, this expectation did not materialize, as the lack of a residual covariation between gF and the ability to learn analogies posed a major model misspecification. The addition of the respective residual term to the model showed that there is a correlation of .47, i.e., there is 22% of shared variance that is not accounted for by WM.

Discussion

The present study focused on the role of executive functions (EF) and fluid intelligence (gF) in explaining the ability of learning to reason analogically. Analogical reasoning is crucial for general ability and that is why it is important to identify the factors that drive its development on a small scale, i.e., within a specific learning situation. Since any higher-order type of learning is by definition a complex process that requires dealing with novelty, there is an inherent need for a system of regulatory processes, frequently labeled as EF. Because the structure of interrelationships between individual EF is highly developmentally specific (Brydges, Fox, Reid, & Anderson, 2014), the first objective was to test whether the proposed theoretical structure matched the empirical relations observed in the data. Based on the testing of a structural equation model, we found that a four-factor structure fitted the data. However, the magnitudes of the relationships suggest that fluency and shifting

are far from being truly diverse and independent from the attentional control, which alone explained 91%, 84%, and 41% of the variation in shifting, fluency and WM, respectively. Possibly, the development of EF may proceed in a similar manner to the development of specific abilities, conforming to Spearman's law of diminishing returns (Spearman, 1927), where the age-related increase in general mental ability (g) leads to the differentiation of cognitive functioning into a system of specific abilities. With regard to executive functioning, it may be the attentional control that drives the development, as postulated by several theories of executive functioning (Barkley, 1997; Pennington & Ozonoff, 1996). On the other hand, functional deficits in attentional control may hamper the development of other, hierarchically superordinate mental functions, triggering a cascade of behavioral deficits (Knight & Grabowecky, 1995). In the present study, the data indicate that there is no relationship between other EF once the effect of attentional control is accounted for. At a given age, executive functioning can thus still be excessively dependent on a finite, shared pool of attentional resources (Roberts & Pennington, 1996), as predicted by the second hypothesis.

However, with respect to the explanatory power of EF, it was not the attentional control, but WM which proved to have a direct predictive effect on gF and on the ability to learn analogies (cf. Buehner, Krumm, & Pick, 2005). The link between the attentional control and the given outcome variables is thus fully mediated by WM (see Ropovik, 2014), indicating a simplex structure (Jöreskog, 1970) where a higher level process is regressed on the processes residing one complexity level below. Likewise, WM was found to act as a

mediator of the speed-gF relationship, where most of the effect of the age-related improvement in WM on gF was itself attributable to the effect of the increase in speed on WM (Demetriou et al., 2014). Such a hierarchical structure is a necessary prerequisite of a causal interpretation defining WM as a subsystem of the ability to learn analogies as well as gF and, in turn, considering attentional control to be involved in the functioning of WM. At the same time, this inter-individual pattern of relationships complies with the development of executive functioning, which has been shown to develop sequentially, from lower-order functions like response inhibition to the higher order ones like shifting or planning (Klenberg et al., 2001).

In line with previous studies where WM was found to have a very strong predictive power with respect to fluid mental ability in younger children as well as adolescents (Brydges et al., 2012; Friedman et al., 2006), the results of the present study support the notion that the ability to maintain and process active representations of information in the presence of interference (Baddeley & Hitch, 1994) is an essential precondition for effective learning in the domain of analogical reasoning as well as the primary constituent of the fluid ability as such. Yet it remains unsettled, which component of WM is primarily responsible for this relationship – whether it is, e.g., the short-term storage (Colom, Abad, Quiroga, Shih, & Flores-Mendoza, 2008) or the central executive (Kane, Conway, Hambrick, & Engle, 2007).

In general, the results of this study support the notion that WM is central to complex cognition as such (Baddeley, 2000). However, despite the observed large effect sizes, WM was not able to fully account for the relationship between gF and the ability to learn analogies, indicating the presence of a factor distinct from WM. WM is thus not the only factor standing behind the relationship between these two aspects of fluid mental ability (i.e., static and dynamic), but within the current model, it was not possible to address the identity of that factor by means of relevant empirical evidence.

Limitations

The present study has some limitations that deserve mention. First, given the dynamic development of executive functions, which is marked by substantial qualitative changes (Anderson, 2002; Huizinga, Dolan, & van der Molen, 2006), the presented conclusions apply only to the general population of children within the given narrow age range. The structure of EF changes over the course of development, particularly as they are recruited for complex tasks (Best, Miller, & Jones, 2009). In older children, executive functioning can be expected to exhibit a rather modular character, where the EF are still related, but far more diverse (Friedman et al., 2006; Miyake et al., 2000), consequently affecting the role of single EF in explaining the ability to learn analogies. Second, as in most of the EF research, the inherently low loadings of EF measures, caused by the over-representation of construct-irrelevant variance (Rabbitt, 1997; Ropovik et al., 2015), are, from a psychometric point of view, a cause for concern. Even with multiple indicators varying in non-executive aspects as much as possible, it is difficult to provide clear-cut formal evidence regarding the exact identity of the measured latent construct and the interpretation of empirical evidence rests on inferential grounds to some extent. Third, the

good fit of the proposed model to the data does not rule out that there may be other models fitting the data equally well and the structural equation modeling is best seen as a primarily disconfirmatory technique that aims to formally reject ill-fitting models (Bollen, 1989). Last, despite a rather complex nomological network (as operationally defined by the tested structural equation model), the cross-sectional nature of the present research design was able to provide only the necessary but not sufficient empirical evidence for any causal interpretations. A more complex model would be needed to control for the effects of possibly confounding variables (e.g., mental speed, short-term storage).

References

Anderson, P. (2002). Assessment and development of executive function (EF) during childhood. *Child Neuropsychology*, *8*, 71-82. http://doi.org/10.1076/chin.8.2.71.8724

Antonakis, J., Bendahan, S., Jacquart, P., & Lalive, R. (2010). On making causal claims: A review and recommendations. *The Leadership Quarterly*, *21*(6), 1086-1120. http://doi.org/10.1016/j.leaqua.2010.10.010

Arbuckle, J. L. (2013). *IBM SPSS Amos 22 User's Guide*. Armonk, NY: IBM.

Baddeley, A. (2000). The episodic buffer: A new component of working memory? *Trends in Cognitive Sciences*, *4*(11), 417-423. http://doi.org/10.1016/S1364-6613(00)01538-2

Baddeley, A. D., & Hitch, G. J. (1994). Developments in the concept of working memory. *Neuropsychology*, *8*(4), 485-493. http://doi.org/10.1037/0894-4105.8.4.485

Barkley, R. A. (1997). Behavioral inhibition, sustained attention, and executive functions: Constructing a unifying theory of ADHD. *Psychological Bulletin*, *121*(1), 65-94. http://doi.org/10.1037/0033-2909.121.1.65

Best, J. R., Miller, P. H., & Jones, L. L. (2009). Executive functions after age 5: Changes and correlates. *Developmental Review*, *29*(3), 180-200. http://doi.org/10.1016/j.dr.2009.05.002

Bollen, K. A. (1989). *Structural equations with latent variables*. NY: Wiley.

Bollen, K. A., & Stine, R. A. (1992). Bootstrapping goodness-of-fit measures in structural equation models. *Sociological Methods & Research*, *21*(2), 205-229. http://doi.org/10.1177/0049124192021002004

Bransford, J. D., Franks, J. J., Yve, N. J., & Sherwood, R. D. (1989). New approaches to instruction: Because wisdom can't be told. In S. Vosniadou & A. Ortony (Eds.), *Similarity and analogical reasoning* (pp. 470-497). NY: Cambridge University Press. http://doi.org/10.1017/CBO9780511529863.022

Brydges, C. R., Fox, A. M., Reid, C. L., & Anderson, M. (2014). The differentiation of executive functions in middle and late childhood: A longitudinal latent-variable analysis. *Intelligence*, *47*, 34-43. http://doi.org/10.1016/j.intell.2014.08.010

Brydges, C. R., Reid, C. L., Fox, A. M., & Anderson, M. (2012). A unitary executive function predicts intelligence in children. *Intelligence*, *40*, 458-469. http://doi.org/10.1016/j.intell.2012.05.006

Buehner, M., Krumm, S., & Pick, M. (2005). Reasoning = working memory [not equal to] attention. *Intelligence*, *33*(3), 251-272. http://doi.org/10.1016/j.intell.2005.01.002

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). NY: Guilford Press.

Colom, R., Abad, F. J., Quiroga, Á. M., Shih, P. C., & Flores-Mendoza, C. (2008). Working memory and intelligence are highly related constructs, but why? *Intelligence*, *36*(6), 584-606. http://doi.org/10.1016/j.intell.2008.01.002

Cowan, N., Morey, C. C., Chen, Z., & Bunting, M. (2007). What do estimates of working memory capacity tell us? In N. Osaka & L. R. H (Eds.), *The Cognitive neuroscience of working memory* (pp. 43-58). NY: Oxford University Press. http://doi.org/10.1093/acprof:oso/9780198570394.003.0003

Delis, D. C., Kaplan, E., & Kramer, J. H. (2001). *Delis-Kaplan executive function system (D-KEFS)*. San Antonio: The Psychological Corporation.

Demetriou, A., Spanoudis, G., Shayer, M., van der Ven, S., Brydges, C. R., Kroesbergen, E., … & Swanson, H. L. (2014). Relations between speed,

working memory, and intelligence from preschool to adulthood: Structural equation modeling of 14 studies. *Intelligence*, *46*, 107-121. http://doi.org/10.1016/j.intell.2014.05.013

Diamond, A. (2013). Executive functions. *Annual Review of Psychology*, *64*(1), 135-168. http://doi.org/10.1146/annurev-psych-113011-143750

Engle, R. W., Tuholski, S. W., Laughlin, J. E., & Conway, A. R. A. (1999). Working memory, short-term memory, and general fluid intelligence: A latent-variable approach. *Journal of Experimental Psychology: General*, *128*(3), 309-331. http://doi.org/10.1037/0096-3445.128.3.309

Friedman, N. P., Miyake, A., Corley, R. P., Young, S. E., Defries, J. C., & Hewitt, J. K. (2006). Not all executive functions are related to intelligence. *Psychological Science*, *17*(2), 172-179. http://doi.org/10.1111/j.1467-9280.2006.01681.x

Goswami, U. (1992). *Analogical reasoning in children*. Hove, UK: Lawrence Erlbaum

Hayduk, L. (2014). Seeing perfectly fitting factor models that are causally misspecified: Understanding that close-fitting models can be worse. *Educational and Psychological Measurement*, *74*(6), 905-926. http://doi.org/10.1177/0013164414527449

Hu, L., & Bentler, P. M. (1999). Cutoff criteria for fit indexes in covariance structure analysis: Conventional criteria versus new alternatives. *Structural Equation Modeling: A Multidisciplinary Journal*, *6*(1), 1-55. http://doi.org/10.1080/10705519909540118

Huizinga, M., Dolan, C. V, & van der Molen, M. W. (2006). Age-related change in executive function: Developmental trends and a latent variable analysis. *Neuropsychologia*, *44*(11), 2017-2036. http://doi.org/10.1016/j.neuropsychologia.2006.01.010

Chen, Z., Sanchez, R. P., & Campbell, T. (1997). From beyond to within their grasp: The rudiments of analogical problem solving in 10- and 13-month-olds. *Developmental Psychology*, *33*(5), 790-801. http://doi.org/10.1037/0012-1649.33.5.790

Cho, S., Holyoak, K. J., & Cannon, T. D. (2007). Analogical reasoning in working memory: Resources shared among relational integration, interference resolution, and maintenance. *Memory & Cognition*, *35*(6), 1445-1455. http://doi.org/10.3758/bf03193614

Jöreskog, K. G. (1970). Estimation and testing of Simplex models. *British Journal of Mathematical and Statistical Psychology*, *23*(2), 121-145.

http://doi.org/10.1111/j.2044-8317.1970.tb00439.x

Kane, M. J., Conway, A. R. A., Hambrick, D. Z., & Engle, R. W. (2007). Variation in working memory capacity as variation in executive attention and control. In A. R. A. Conway, A. Jarrold, M. J. Kane, A. Miyake, & J. N. Towse (Eds.), *Variation in working memory* (pp. 21-48). Oxford: Oxford University Press. http://doi.org/10.1093/acprof:oso/9780195168648.003.0002

Klenberg, L., Korkman, M., & Lahti-Nuuttila, P. (2001). Differential development of attention and executive functions in 3- to 12-year-old Finnish children. *Developmental Neuropsychology*, *20*(1), 407-428. http://doi.org/10.1207/s15326942dn2001_6

Knight, R. T., & Grabowecky, M. (1995). Escape from linear time: Prefrontal cortex and conscious experience. In M. S. Gazzaniga (Ed.), *The cognitive neurosciences* (pp. 1357-1371). Cambridge: MIT Press.

Kroger, J. K. (2002). Recruitment of anterior dorsolateral prefrontal cortex in human reasoning: A parametric study of relational complexity. *Cerebral Cortex*, *12*(5), 477-485. http://doi.org/10.1093/cercor/12.5.477

Kyllonen, P. C., & Christal, R. E. (1990). Reasoning ability is (little more than) working-memory capacity?! *Intelligence*, *14*(4), 389-433. http://doi.org/10.1016/s0160-2896(05)80012-1

Lehto, J. E., Juujärvi, P., Kooistra, L., & Pulkkinen, L. (2003). Dimensions of executive functioning: Evidence from children. *British Journal of Developmental Psychology*, *21*(1), 59-80. http://doi.org/10.1348/026151003321164627

MacCallum, R. C., Browne, M. W., & Sugawara, H. M. (1996). Power analysis and determination of sample size for covariance structure modeling. *Psychological Methods*, *1*(2), 130-149. http://doi.org/10.1037/1082-989x.1.2.130

MacCallum, R. C., Roznowski, M., & Necowitz, L. B. (1992). Model modifications in covariance structure analysis: The problem of capitalization on chance. *Psychological Bulletin*, *111*(3), 490-504. http://doi.org/10.1037/0033-2909.111.3.490

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., Howerter, A., & Wager, T. D. (2000). The unity and diversity of executive functions and their contributions to complex "Frontal Lobe" tasks: A latent variable analysis. *Cognitive Psychology*, *41*(1), 49-100. http://doi.org/10.1006/cogp.1999.0734

Oberauer, K., Süß, H.-M., Wilhelm, O., & Wittmann, W. W. (2008). Which working memory functions predict intelligence? *Intelligence*, *36*(6), 641-652. http://doi.org/10.1016/j.intell.2008.01.007

Pennington, B. F., & Ozonoff, S. (1996). Executive functions and developmental psychopathology. *Journal of Child Psychology & Psychiatry*, *37*(1), 51-87. http://doi.org/10.1111/j.1469-7610.1996.tb01380.x

Rabbitt, P. (1997). Introduction: Methodologies and models in the study of executive function. In P. Rabbitt (Ed.), *Methodology of frontal and executive function* (pp. 1-38). East Sussex, UK: Psychology Press.

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*. http://doi.org/10.1016/S0885-2014(98)90003-X

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://doi.org/10.1016/j.jecp.2006.02.002

Roberts, R. J., & Pennington, B. F. (1996). An interactive framework for examining prefrontal cognitive processes. *Developmental Neuropsychology*, *12*(1), 105-126. http://doi.org/10.1080/87565649609540642

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

Ropovik, I. (2015). A cautionary note on testing latent variable models. *Frontiers in Psychology*, *6*, 1715. http://doi.org/10.3389/fpsyg.2015.01715

Ropovik, I., Bobakova, M., Ferjencik, J., Filickova, M., Kovalcikova, I., & Slavkovska, M. (2015). The utility of set-loss error scores in the general population. *Psychological Assessment*, *27*(3), 755-762. http://doi.org/10.1037/pas0000106

Ruef, M., Furman, A., & Muńoz-Sandoval, A. (2003). *Woodcock-Johnson: International editions. Administration manual: Slovak edition*. Nashville: The Woodcock-Muńoz Foundation.

Spearman, C. E. (1927). *The abilities of man*. London: Macmillan.

Stevenson, C. E., Bergwerff, C. E., Heiser, W. J., & Resing, W. (2014). Working memory and dynamic measures of analogical reasoning as predictors of children's math and reading achievement. *Infant and Child Development*, *23*(1), 51-66. http://doi.org/10.1002/icd.1833

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. http://doi.org/10.1016/j.intell.2013.01.003

Tabachnick, B. G., & Fidell, L. S. (2007). *Using multivariate statistics*. Boston: Pearson/Allyn & Bacon.

Vygotsky, L. S. (1987). *Thinking and speech*. NY: Plenum Press.

Wiebe, S. A., Espy, K. A., & Charak, D. (2008). Using confirmatory factor analysis to understand executive control in preschool children: I. Latent structure. *Developmental Psychology*, *44*(2), 575-587. http://doi.org/10.1037/0012-1649.44.2.575