UC Merced

Proceedings of the Annual Meeting of the Cognitive Science Society

Title

The Role of Executive Functions for Structure-Mapping in Mathematics

Permalink

https://escholarship.org/uc/item/02j5t4pf

Journal

Proceedings of the Annual Meeting of the Cognitive Science Society, 37(0)

Authors

Begolli, Kreshnik Nasi

Richland, Lindsey Emgle

Jaeggi, Susanne

Publication Date

2015

Peer reviewed

The Role of Executive Functions for Structure-Mapping in Mathematics

Kreshnik Nasi Begolli[1] (kbegolli@uci.edu), Lindsey Engle Richland[2] (lrichland@uchicago.edu), and Susanne Jaeggi[1] (smjaeggi@uci.edu)

[1]School of Education, 3200 Education, Irvine, CA 92697 USA
[2]Department of Comparative Human Development, 5730 S. Woodlawn Ave., Chicago, IL 60637

Abstract

Comparing analogs is a key recommendation in mathematics instruction, but successful structure-mapping may impose high demands on children's executive functions (EF). We examine the role of individual differences in EF resources on learning from an everyday mathematics video-lesson placing a particular strain on children's cognitive resources: comparing three analogs presented sequentially. Specifically, we examine the separate contributions of working memory (WM) and inhibitory control (IC) on successful schema-formation. Overall, WM and IC explained distinct variance for predicting improvements in procedural knowledge, procedural flexibility, and conceptual knowledge after a 1-week delay. WM & IC are less predictive at immediate post-test, suggesting that these functions are not simply correlated with mathematics skill, but may be particularly important in the process of structure-mapping for durable schema-formation. These results inform the literature on both analogy and mathematics learning, extending previous findings implicating EFs as key for successful structure-mapping to an ecologically valid learning context.

Keywords: analogy; comparison; mathematics education; video stimulus; misconception; executive function, inhibitory control, working memory

Identifying contrasts and similarities between multiple representations is central for developing conceptual knowledge in mathematics (see NRC, 2001) and for inducing conceptual change (Vosniadou, Vamvakoussi, and Skopelitti, 2008). However, this engages complex cognitive processes that place a burden on reasoners' executive functions (EFs). Learners use EFs to perform relational structure-mapping: represent systems of relationships, align and map these systems to each other, and draw inferences based on the alignments (and misalignments) for successful schema formation (see Gentner, 1983; Gick & Holyoak, 1983; Morrison et al., 2011).

Orchestrating structure-mapping opportunities in classroom lessons that lead to successful schema formation is not straightforward, particularly because reasoners regularly fail to notice the utility of aligning and mapping two or more available relational structures (Gick & Holyoak, 1983; Ross, 1989), often leading to misconceptions (Zook, 1991).

A common instructional recommendation to help students confront misconceptions is to directly contrast them with valid relational structures, (i.e. in this case, solution strategies). At the same time, engaging with misconceptions without fully encoding the higher order relation between that misconception and the correct analog may also lead to reification of these intuitions (Begolli & Richland, under review).

In this study we explore the hypothesis that children's success in overcoming misconceptions through comparisons with correct analogs may vary based on limitations in children's developing EF (see Waltz et al., 2000). Because misconceptions are often deeply embedded in intuitive beliefs, drawing a higher order relation between a misconception and a correct analog to form a valid schema is highly effortful and requires a combination of executive functions. WM is argued to be necessary for representing systems of objects (e.g. steps to solution strategies) and re-representing these systems of relationships in order to align and map their structures. Successful mapping and alignment requires flexible switching between these systems of relations to attend to relevant elements within each system and inhibit irrelevant elements to identify meaningful similarities and differences, in order to derive conceptual/schematic inferences from this structure-mapping exercise and better inform future problem solving (see Morrison et al., 2011). Thus, limitations of EFs – working memory, task switching, and inhibition throughout this reasoning process could explain failures in schema formation through structure-mapping.

In a previous experiment, Begolli & Richland (2013), presented students with a common misconception followed by two correct solution strategies and examined whether presenting analogs either simultaneously, sequentially, or only verbally would support structure-mapping in a mathematics lesson based on instructional analogy. They found that students' schema formation was best supported when analogs were visible throughout the structure-mapping. However, sequential presentations of analogs led to the lowest performance suggesting that object-level encoding of misconceptions interfered with schema formation, perhaps due to limitations of EFs (Begolli & Richland, 2013). Sequential presentation of analogs may place a greater strain on EF resources, potentially revealing EF mechanisms responsible for structure-mapping failures.

This study examines correlations between schema formation from sequential presentation of analogs (as in Begolli & Richland, 2013) and individual difference measures of EF – particularly working memory processes (WM; short-term and domain general WM) and inhibitory control processes (IC; response inhibition and task switching). Working memory is likely to facilitate the manipulation of relational systems while holding them in mind and IC is hypothesized to decrease distractional

1

elements within these systems, enable disattention to an intuitive misconception, and aid in switching between relations to derive appropriate schemas.

This work has the potential to contribute to both the theoretical understanding of the role of EFs in successful structure-mapping within the ecologically valid context of a classroom as well as practical implications for designing technology and instruction.

Study: EF and Instructional Analogy

Method

Participants. Participants were 107 5th graders (44 girls) drawn from a school with high socioeconomic status, but 16 students either missed a test or a cognitive measure or both due to absences. Ten additional participants were dropped from analyses because their pretest scores for procedural & conceptual knowledge were at ceiling (100%). The maximum number of participants at each test point and cognitive measure was included in the analyses (n = 88-81).

Design & Procedure. All participants followed the same procedure. Day 1: pretest and individual difference measures of EF. Day 2: (2 days later), interactive instructional video as the intervention, followed by an immediate posttest. Day 3 (1 week later): delayed posttest and an additional EF measure.

Instructional Stimuli. Because the study takes ecological validity and the complexity of everyday classrooms as serious constraints, a novel methodology was used to derive rigorous data that incorporates the complexities of situated cognition. The stimuli derived from videotapes of a teacher in her classroom, teaching a lesson co-designed with the research team (for more detail see Begolli & Richland, 2013; Shimizu, 2003). The teacher guides students to draw connections between three solution strategies to a ratio problem (Figure 1 left): subtraction (incorrect), least common multiple and division (both correct).

Ratio was chosen as an instructional topic for two reasons: (a) it is part of the common core standards for elementary mathematics instruction and (b) previous research has shown that ratio problems prompt diverse systematic student responses, useful for charting trajectories of reasoning change across the study.

Mathematics Assessment. The assessment was designed to assess schema formation and generalization, adapted from Begolli & Richland (2013). Mathematically, the assessment included procedural knowledge, procedural flexibility, and conceptual knowledge constructs. The constructs were conceptually derived from Rittle-Johnson and Star (2009), and adapted to core concepts and procedures underlying ratio problems. The three constructs were measured through

multiple items, averaged to derive an overall composite score for that particular construct.

Procedural Knowledge, Procedural Flexibility, and Conceptual Knowledge. The procedural knowledge (PK) construct measured whether students could produce solutions of familiar and near transfer problems. The procedural flexibility (PF) construct measured: (a) students' adaptive production of solution methods, (b) their ability to identify the most efficient strategy, and (c) students' ability to identify a novel solution method which was related to a taught strategy. The conceptual knowledge (CK) construct was designed to probe into students' explicit and implicit knowledge of ratio (see Figure 1).

Misconception Usage Score. Misconceptions are mistakes that students make, which obstruct learning (Smith, diSessa, Roschelle, 1994). Based on a published lesson (Shimizu, 2003), pilot and pretest data, a solution involving subtraction was expected to be the most common misconception (CM) participants would bring to the study. For example, in the problem shown in Figure 1, students would subtract total shots tried 12 from total shots made 20 and compare who missed more shots. The CM score was the average number of times a student produced or chose subtraction on such problems. This score assessed students' ability to overcome their misconceptions about how to solve rate and ratio problems as well as the conditions under which students confirm invalid biases.

Figure 1. Procedural/Procedural flexibility problem (left) used in the video-lesson and assessments, and conceptual problem used in assessments (right). For procedural flexibility students were told to solve a problem similar to the one on the left using two different strategies.

Measures of Executive Functions. EF measures were administered to examine relations between individual differences in students' processing resources and learning from the video-lesson.

Forward and Backwards Digit Span (Administered Day 1) The Forward Digit Span (FDS; repeat numbers in the same order) is a measure of short-term memory (storage), whereas the Backward Digit Span (BDS; repeat numbers in reverse order) is a measure of domain general WM processes (storage + processing). The maximum set of numbers recalled twice correctly was used as a dependent measure on both the FDS and the BDS.

Hearts and Flowers. (Day 1) The Hearts and Flowers task (H&F) is a version of the Dots task taken from the Directional Stroop Battery used to assess EF (adapted from Wright & Diamond 2014). This was administered on day 1.

Students were presented with either hearts (congruent) or flowers (incongruent) on each trial (Figure 2). For

2

incongruent trials, the correct response is aligned with students' natural inclination – "press the button on the same side (left or right) as the heart." For incongruent trials, the correct response goes against what comes naturally – "press the button on the opposite side (left or right) of the flower." Trials are presented in 3 phases. Phase 1 – congruent trials only, phase 2 – incongruent trials only, phase 3 – mixed trials presented randomly.

To perform this task students are expected to hold each task in mind (short-term memory), switch between tasks to choose the right answer (task switching), and inhibit their pre-potent response (see Wright and Diamond, 2014). The dependent measure was the difference in time it took to respond to a trial when participants had to change the rule versus a trial when participants did not have to change the rule to respond – known as switch cost response time.

Stop-Signal Task (Administered Day 3). The Stop-Signal task (SST) measured participants' response inhibition. Students are presented with a fish for 850ms (go stimulus) or a fish followed by a manta ray (stop-signal, occurring on 40% of the trials). Students were instructed to press a button ("A" or "L") to send the fish home (within 850ms) unless the Manta Ray appeared, in which case they had to withhold from pressing any buttons. The sooner the Stop-Signal appears after the go signal, the easier it is to inhibit a response – this temporal difference is known as the Stop-signal Delay (SSD). SSDs are initially short but are increased following accurate trials. Final SSD length was used as a dependent measure (Bissett and Logan, 2012)

Analyses. Executive Functions share *commonalities*, but also have *diverse* functions, for controlling thought and behavior (Miyake et al., 2000). To understand whether the contribution of each cognitive measure was separable or unitary we conducted a principal factor analysis with a varimax rotation on all measures (FDS & BDS, H&F, and SST; Table 1). Combining measures also reduces task specific variance and allows examination on a construct level, rather than on an individual task level. The theoretical expectation was to derive two distinct factors sharing common variance. A WM factor to account for the common contribution of short-term and domain general working memory processes (comprised of the FDS & BDS) and an IC factor accounting for the common contribution of response inhibition and task switching processes (comprised of the H&F and SST). The results of the factor analyses confirmed these predictions with both factors displaying similar loadings which explained 65.1% of the total variance. The factor scores for the WM and IC factors were converted into z-scores for subsequent analyses.

Table 1. Factor Loadings and Descriptives

N = 96	WM Factor	IC Factor	Mean	SD
FDS	**0.83**	0.01	6.05	1.11
BDS	**0.78**	0.06	5.36	1.08
H&F	0.15	**0.77**	110*	169
SSD	-0.07	**0.82**	284	164
% of Variance	33.2%	31.8%		

*Median RT used & reported here. H&F and SSD are in ms

To examine the contribution of broader WM and IC as well as to unpack the contribution of each cognitive process, we conducted separate regressions on each mathematics construct (PK, PF, CK, and CM) for three models at pretest, immediate, and delayed test, summarized in Table 2.

Table 2. Regression models conducted in analyses. A separate regression was conducted for each mathematics construct

Dependent Measure	Model 1	Model 2	Model 3
	Math Construct	Math Construct	Math Construct
	WM Factor	FDS	H&F
Indicators	IC factor	BDS	SST
	Pretest score*	Pretest score*	Pretest score*

*Only utilized at immediate and delayed posttests.

The first model examines the role of WM & Inhibition as key processes in EF on each mathematical construct separately. Model 2 unpacks the role of WM by examining the individual contribution of short-term (FDS) and domain general WM processes (BDS). Model 3 unpacks the role of IC by examining the individual contribution of task switching (H&F) and response inhibition (SST).

Results

Tables 1 and 3 summarize the mean scores of the cognitive measures and mathematical constructs. Regression results with beta values for all Models on each mathematical construct are summarized in Table 4.

Irrespective of cognitive ability, students improved from pretest to immediate and delayed posttest on PK, PF, and CK ($F > 10$, $p < .000$), but no overall difference on how much students' used the misconception ($F = 1.04$, $p = .23$). But differences in students' EF may reflect distinct patterns in their math outcomes. The remaining results will be discussed by presenting the relationship between WM and each math construct from Model 1, then we discuss Model 2 to unpack the contribution of each component within WM, *Short-Term Memory* (FDS) and *domain general WM* (BDS) on each construct. Similarly, we discuss the relationship between IC and each math construct from Model 1, and then in Model 3, unpack the contribution of *Response Inhibition* (SST) and *Task Switching* (H&F) within IC. Effect sizes for each component ranged from small ($\eta^2 = .02$) to moderate ($\eta^2 = .14$). Model 2 and Model 3 analyses were exploratory and thus there was no family-wise error correction, however the results should be interpreted with caution.

Table 3. Mean scores per construct, SD in parenthesis.

	Pretest ($N = 87$)	Immediate ($N = 91$)	Delayed ($N = 92$)
Procedural	22% (0.25)	48% (0.36)	45% (0.37)
Flexibility	11% (0.12)	27% (0.22)	24% (0.22)
Conceptual	32% (0.27)	43% (0.28)	45% (0.31)
Misconcept.	25% (0.24)	20% (0.23)	23% (0.26)

Working Memory. Students' WM ability does not seem to predict pretest performance, though when unpacking the

3

WM factor, BDS performance was positively related with higher uses of the common misconception (η^2 =.06).

At immediate test, overall WM ability was positively related with conceptual knowledge performance, which seems to be largely driven by advantages in students' FDS scores (η^2 =.09).

At delayed test, students' with higher WM factor scores had overall higher outcomes in procedural knowledge (η^2 =.09), procedural flexibility (η^2 =.05), and conceptual knowledge (η^2 =.08). When looking at the individual contribution of each WM component, only students with higher FDS scores had higher scores in procedural knowledge (η^2 =.05). While the relationship between students with higher BDS scores and conceptual knowledge scores was not significant, it suggested a positive trend (p = .052; η^2 =.05).

Table 2. Beta values from regression models described in Table 1. Pretest beta values were always significant, p <.05; not shown here

	Pretest		Immediate		Delayed	
Model 1	WM	IC	WM	IC	WM	IC
Procedural	4.08	-0.02	2.65	5.70†	**8.49****	**10.76*****
Flexibility	0.70	-0.50	3.68†	3.98†	**4.58***	**5.30***
Conceptual	0.65	**7.18***	**6.20***	3.20	**6.99***	**6.68***
Misconcep.	2.38	3.27	-3.76	**-5.26***	-3.58	**-6.91***
Model 2	FDS	BDS	FDS	BDS	FDS	BDS
Procedural	3.01	2.14	2.34	0.64	**7.18***	3.04
Flexibility	1.18	-0.46	1.55	3.41	2.12	4.15
Conceptual	0.82	-0.63	**8.11****	0.03	2.96	5.70†
Misconcep.	-3.00	**6.32***	-3.91	-0.62	-2.62	-1.10
Model 3	SST	H&F	SST	H&F	SST	H&F
Procedural	-3.92	3.85	4.66	3.23	**8.01***	**6.74***
Flexibility	-1.07	0.62	1.94	3.38	**5.61***	1.47
Conceptual	5.26†	4.29	2.08	2.84	3.68	5.43†
Misconcep.	4.73†	-0.62	-1.86	**-5.14***	-1.23	**-7.41****

† $p < .10$, * $p < .05$, ** $p < .01$, *** $p < .001$

Inhibitory Control. At pretest, students' with higher overall IC scores reflect an advantage only in their conceptual knowledge performance (η^2 =.07).

At immediate test, students' with better IC scores use the misconception less (η^2 =.05), which is positively related to their task switching performance measured by the H&F task (η^2 =.06).

At delayed test, students with higher scores in IC display an advantage in their procedural knowledge (η^2 =.14), procedural flexibility (η^2 =.07), conceptual knowledge (η^2 =.07) constructs and a reduction in their use of misconceptions (η^2 =.07). In these cases, students with higher SST scores are better in procedural knowledge (η^2 =.07) and procedural flexibility (η^2 =.07) . Students with higher H&F scores are better in procedural knowledge (η^2 =.06) and use the misconception less (η^2 =.08). The relationship between IC and conceptual knowledge could be driven by students' task switching performance, though this relationship is not significant (p = .057; η^2 =.05).

To help interpret the data from another perspective and for illustration purposes, we divided students based on their

WM & IC scores into high (top 25%), medium (middle 50%), & low (bottom 25%) performers (Figure 3).
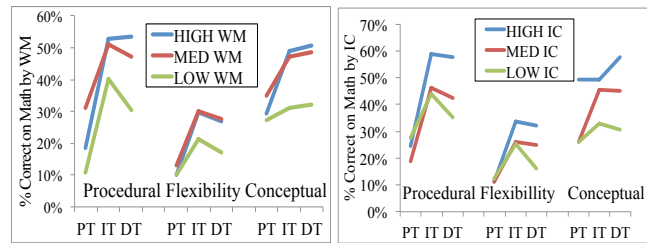

Figure 3. Mean scores for Pretest (PT), Immediate (IT) and Delayed (DT) tests by WM (left) and IC (right) score.

The regression results suggest a continuous progression between low, medium, and high performers, such that students with a 1-point advantage in WM or IC score have advantages on their mathematics performance ranging from roughly 20%-29% higher than the overall mean.

A qualitative examination of the WM data suggests that this effect is driven by a difference between low and medium/high WM performers, with the largest differences on the conceptual understanding measures. Procedural knowledge has been proposed to be a preliminary step for attaining conceptual knowledge, though both reinforce each other iteratively (Rittle-Johnson et al., 2001). It appears that low WM students show some improvements in their procedural and flexible knowledge at immediate test, but these gains decrease by delayed test, perhaps reflecting a level of procedural knowledge that is insufficient for retention, nor for attaining a broader schema for ratio – reflected in conceptual understanding measures. This perspective reinforces the role of domain general WM processes (in contrast to short-term processes) as critical for durable schema formation, but may also indicate that there is a certain threshold for WM ability required. Thus students must have adequate WM for schema formation, but performance may not be affected if their WM ability passes this threshold.

In contrast, a qualitative examination of the IC measures lend themselves towards interpreting a more continuous relationship between IC ability and students' mathematics outcomes. Students with high and medium IC scores use the misconception more before the lesson (pretest M = 28%) than after the lesson (delayed M= 19%), while medium IC students remain about the same (pretest M = 28%, delayed M = 25%). An inverse relationship seems true for low IC students (pretest M = 17%, delayed M = 27%). Further research is needed in order to clarify the interpretations stemming from these exploratory perspectives.

In sum, both WM and IC predict procedural, flexible, and conceptual knowledge at delayed test and IC also predicts a reduction in misconceptions. These effects are less apparent at immediate test, suggesting that WM and IC may be particularly important for gaining a deeper, more schematic understanding of concepts, which in turn may promote flexible knowledge and retention of procedures. Negative

correlations between IC and the misconception implies that inhibitory processes may be required to reduce misconceptions.

Discussion

This study clarifies the contribution of EF abilities for schema formation of mathematics concepts through instructional analogies. Many studies have examined the relationship between EF and broader mathematics achievement (St Clair-Thompson & Gathercole, 2006), but there is little work done investigating the specific role of EF on *learning* mathematics through structure-mapping. Unlike previous accounts that have shown relationships between EFs and structure-mapping (Zelazo et al., 2003; Waltz et al., 2000; Morrison et al., 2011; Krawczyk et al., 2008; Richland & Burchinal, 2012) the pretest, intervention, posttest design in combination with the EF measures gives insight into the role of specific EFs throughout the trajectory of schema-formation by structure-mapping.

In addition, the WM & IC data align with current views that WM & IC are separate processes within EF, each explaining distinct variance (Miyake et al., 2000). These data reveal that within WM, the short-term and general working memory processes share commonalities, but each also accounts for distinct variance in an everyday analogical learning context. Similarly, within IC, response inhibition and task switching share commonalities, but also account for distinct variance in learning from analogy.

The results reveal that broader WM and IC processes predict learning in this instructional context. Both WM and IC predicted the retention of procedural knowledge, procedural flexibility, and conceptual knowledge, and IC also predicted the reduction of students' use of the misconception, reflected by delayed test results. EF resources (WM and IC) may matter most for durable schema formation, while their effect may be less evident for short-term learning, as evidenced by their more limited prediction of performance at immediate test.

In the short term, it seems that the relationship between WM and conceptual knowledge is largely influenced by short-term memory processes, whereas the relationship between students' use of misconceptions and their IC ability seems to be driven by students' task switching performance. A possible explanation is that in the structure mapping process, short-term memory processes facilitate the representation of systems of relations, whereas task switching processes (which include inhibition) help reasoners attend to structural dissimilarities between these systems. However, at immediate test, it may be hard to distinguish between recency effects/object-level encoding and successful schema formation, which could also obstruct from understanding the role of WM and IC (and/or individual functions within WM and IC) in the long term. Thus, examining delayed test results provides better data on the role of WM and IC on successful structure-mapping.

A closer inspection of delayed tests results suggests that WM and IC components have the most predictive power when considered in tandem, as their individual contributions wane when considered separately. In terms of WM components, short-term storage seems to be related to only procedural knowledge, whereas general WM seems somewhat related to the attainment of conceptual knowledge, though not significant perhaps due to low power ($p = .052$).

A closer examination of separate IC processes allows for hypothesis generation about the specific EF resources and their relations to learning. While these data should be interpreted with caution, the data patterns suggest that students that are better on the response inhibition task (SST) have higher procedural knowledge and are more flexible with procedures. Further, students who perform better on task switching (H&F) continue to use misconceptions less – and there is somewhat of a relationship between task switching and conceptual knowledge, though only marginally significant ($p = .057$).

In light of delayed test data, it appears that short-term memory processes (FDS) may be important for initial schema-formation and for later recall of the appropriate procedures, whereas general WM processes as measured by the BDS are more important for long-term generalizable knowledge. Perhaps students with better FDS scores had greater resources to represent the systems of relations during the structure-mapping processes, but only those students with better BDS scores were able to re-represent these systems for appropriate alignment and mapping between the source and target relations, leading to a more durable schema.

On the other hand, students who were better at response inhibition (SST) and task switching (H&F) may not notice their advantage immediately, but these processes may be crucial for long-term schema formation. It could be the case that better response inhibition during the structure-mapping process aids students' WM to attend to appropriate representations by reducing interference from competing and inappropriate representations (in this case ratio concepts over subtraction – the common misconception). Thus, leading towards increased procedural knowledge and flexibility. Another interpretation, though not mutually exclusive, is that response inhibition is responsible for reducing competing representations and selecting the correct representation at the time of the assessment 1-week later (though this may also imply reductions in the use of misconception, not reflected in the data).

It appears that task-switching processes operate at a higher level such that at every switching point, response inhibition may be required to select the appropriate task. This process seemed likely to lead towards an increase in conceptual understanding and a reduction of misconceptions. A possible explanation is that in order to identify relations that structurally align in the source and target representations the reasoner has to repeatedly switch between these representations while inhibiting distracting information in order to successfully map their structural relations. Overall, the data from WM and IC measures align

5

with previous neurological and behavioral data, and computational models suggesting a similar role for inhibitory control (e.g. LISA; Morrison et al., 2011; Zelazo et al., 2003; Waltz et al., 2000; Morrison et al., 2011; Krawczyk et al., 2008; Richland & Burchinal, 2012).). Previous behavioral data have suggested that increases in relational complexity within analogs would place a higher demand on children's EF resources (Halford et al., 2002). Also, populations with compromised EF resources (e.g., damaged PF cortex) or strained EFs (students performing dual-tasks during structure mapping; Waltz et al., 2000), and younger children (Richland et al., 2006) are more likely to fail at structure-mapping. Broader EF and IC at 54-months have been found to predict analogical reasoning at age 15 (Richland & Burchinal, 2012). Thus, there is mounting evidence converging on the importance of WM and IC as underpinnings of analogical reasoning.

In sum, in an ecologically valid learning context, our data provide evidence that individual differences in EF may impact whether students successfully benefit from a structure-mapping opportunity comparing a misconception to correct solutions. Teachers wishing to confront students' misconceptions, thus, may be helping students with high EF resources while harming those with low EF resources when sequentially presenting these analogs in their lessons. Simultaneous presentations of analogs may reduce the disparity in schema-formation due to individual differences in EF, but this remains to be tested.

References

Begolli, K.N., & Richland, L.E. (2013) Visual Support for Instructional Analogy: Context Matters, Paper presented at the Cognitive Science Society, Berlin, Germany.

Begolli, K. N. & Richland, L.E. (under review). Teaching Mathematics by Comparison: The Visibility of Analogs Matters.

Bissett, P. G., & Logan, G. D. (2012). Post-stop-signal slowing: Strategies dominate reflexes and implicit learning. *Journal of Experimental Psychology: Human Perception and Performance*, *38*(3), 746.

Gentner, D. (1983). Structure-mapping: A theoretical framework for analogy. *Cognitive Science*, 7, 155-170.

Gick, M.L. & Holyoak, K. J. (1983). Schema induction and analogical transfer. *Cognitive Psychology,15,* 1-38.

Halford, G. S., Andrews, G., Dalton, C., Boag, C., & Zielinski, T. (2002). Young children's performance on the balance scale: The influence of relational complexity. *Journal of Experimental Child Psychology*, *81*(4), 417-445.

Krawczyk, D. C., Morrison, R. G., Viskontas, I., Holyoak, K. J., Chow, T. W., Mendez, M. F., ... & Knowlton, B. J. (2008). Distraction during relational reasoning: the role of prefrontal cortex in interference control. *Neuropsychologia*, *46*(7), 2020-2032.

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.

Morrison, R. G., Doumas, L. A. A., and Richland, L. E. (2011). A computational account of children's analogical reasoning: Balancing inhibitory control in working memory and relational representation. Developmental Science, 14, 516–529.

National Research Council. (2001). *Adding it up: Helping children learn mathematics*.

Richland, L. E., & Burchinal, M. R. (2012). 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.

Rittle-Johnson, B., Siegler, R. S., & Alibali, M. W. (2001). Developing conceptual understanding and procedural skill in mathematics: An iterative process. *Journal of educational psychology*, *93*(2), 346.

Rittle-Johnson, B., and Star, J. R. (2009). Compared to what? The effects of different comparisons on conceptual knowledge and procedural flexibility for equation solving. *Journal of Educational Psychology, 101*, 529–544.

Ross, B.H. (1989). Distinguishing types of superficial similarities: Different effects on the access and use of earlier problems. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 15,*456–468.

Shimizu, Y. (2003). Problem Solving as a Vehicle for Teaching Mathematics: A Japanese Perspective. 205-214. In F. K. Lester Jr. (Ed), *Teaching Mathematics through Problem Solving. Prekindergarten - Grade 6*. National Council of Teachers of Mathematics.

Smith J.P., III, diSessa A. A., and Roschelle J. (1994). Misconceptions Reconceived: A Constructivist Analysis of Knowledge in Transition. *The Journal of Learning Sciences*. 3(2), 115-163.

St Clair-Thompson, H. L., & Gathercole, S. E. (2006). Executive functions and achievements in school: Shifting, updating, inhibition, and working memory. *The quarterly journal of experimental psychology*, *59*(4), 745-759.

Vosniadou, S., Vamvakoussi, X., & Skopeliti, I. (2008). The framework theory approach to the problem of conceptual change. *International handbook of research on conceptual change*, 3-34.

Waltz, J.A., Lau, A., Grewal, S. K., Holyoak, K.J. (2000). The role of working memory in analogical mapping. Memory and Cognition, 28(7):1205-12.

Zelazo, P. D., Müller, U., Frye, D., Marcovitch, S., Argitis, G., Boseovski, J., ... & Carlson, S. M. (2003). The development of executive function in early childhood. *Monographs of the society for research in child development*, i-151.

Zook, K. B. (1991). Effects of analogical processes on learning and misrepresentation. *Educational Psychology Review,* 3, 41-72.

6