Executive Functions and Conceptual Change in Science and Mathematics Learning

Stella Vosniadou (svosniad@phs.uoa.gr)
Department of Philosophy and History of Science, University of Athens
Panepistimioupolis, GR.157 71 Athens – Greece, and
The Flinders University of South Australia

Dimitrios Pnevmantikos (dpnevmat@uowm.gr)
Department of Primary Education, University of Western Macedonia,
GR.531 00, Florina – Greece

Nikos Makris (nmakris@eled.duth.gr)
Department of Primary Education, Democritus University of Thrace,
University Campus, GR.68100 Alexandroupolis – Greece

Kalliopi Ikospentaki (**kikospe@phs.uoa.gr**)
Department of Philosophy and History of Science, University of Athens
Panepistimioupolis, GR.157 71 Athens – Greece

Despina Lepenioti (**des_lepenioti@hotmail.com**)
Department of Philosophy and History of Science, University of Athens
Panepistimioupolis, GR.157 71 Athens – Greece

Anna Chountala (**annachountala@gmail.com**)
Department of Philosophy and History of Science, University of Athens
Panepistimioupolis, GR.157 71 Athens – Greece

Giorgos Kyrianakis (kyrianakis_giorgos@yahoo.gr)
Department of Primary Education, University of Western Macedonia,
GR.531 00, Florina – Greece

Abstract

We investigate the hypothesis that Executive Functions (EFs) are implicated in the learning of science and mathematics by examining the relation between performance in two Science and Mathematics Conceptual Understanding and Conceptual Change (CU&C) tasks, and two Stroop-like Inhibition and Shifting EF tasks, in a group of 69 4th and 6th grader children. The results showed high correlations between accuracy performance in the CU&C and EF tasks even when Intelligence Ability (IA) and Age were partialed out. A path analytic model showed that performance in the CU&C tasks could be explained by performance in the EF and IA tasks, which were positively related to each other. Further analyses showed that accuracy of performance particularly in the CU&C tasks could be predicted by performance in the EF tasks, with high or medium EF scores being a prerequisite for placement in the group of high CU&C achievers.

Keywords: Executive Functions, Conceptual Change, Science Learning and Teaching, Mathematics Learning and Teaching

Executive Functions and Conceptual Change in Science and Mathematics Learning

One of the main reasons why children find it hard to understand many concepts in science and mathematics is because these concepts are incompatible with their prior knowledge and require major conceptual changes to take place. For example, cognitive developmental research has shown that by the time systematic science instruction starts children have already constructed a naive physics, which is based on everyday experience and is very different from currently accepted science (Carey, 2009; Chi, 2008). Learning science requires considerable conceptual changes to take place in this naïve physics, such as the construction of new ontological categories, new representations, and new epistemological understandings. Similarly, the acquisition of more advanced concepts in mathematics also requires fundamental conceptual changes in the initial concept of number, which is based on natural number knowledge (Vosniadou & Verschaffel, 2004). Difficulties with fractions, for example, have been associated with students' whole number knowledge (*i.e.* applying whole number knowledge in situations where this is not appropriate), a phenomenon known as the whole number bias (Ni & Zhou, 2005). Natural numbers have a different symbolic representation than fractions, they are discrete and not dense, and have a different relationship to the unit. Children do not understand the symbolic representation of fractions; they often treat fractions' numerators and denominators as separate whole numbers (Stafylidou & Vosniadou, 2004), and do not understand the density and infinite divisibility of number.

Research in the learning of science and more recently in the learning of mathematics suggests that initial concepts based on everyday experience are not replaced by scientific ones but continue to exist and to influence problem solving even when conceptual change has been achieved (DeWolf & Vosniadou, 2015; Dunbar, Fugelsang & Stein, 2007; Shtulman & Valcarcel, 2012). If it is true that initial concepts continue to exist alongside the more recently acquired scientific concepts and are easier and faster to activate, it then follows that access to scientific concepts may require the inhibition of initial concepts and the continuous shifting between initial and scientific concepts. If this is the case then the learning of science and mathematics would implicate the types of cognitive abilities known as Executive Functions (EFs) (e.g. Miyake et al., 2000).

Executive Functions (EFs) are a set of processes responsible for the regulation and monitoring of complex cognitive tasks that require deliberate planning, impulse control, goal-directed behaviour, and flexible strategy employment (Miyake et al., 2000). Three EFs are considered as particularly important: *inhibition, shifting, and updating. Inhibition* refers to the ability to suppress dominant and automatic responses in favour of a more appropriate goal appropriate response; *shifting* is defined as the ability to switch between sets of tasks or strategies, and *updating* refers to the ability to update the content of working memory by replacing old items with new information. EF abilities develop rapidly during the preschool and school years (Davidson et al., 2006) and are highly associated with academic readiness and math and reading scores (Blair et al., 2007; van de Sluis et al., 2007).

The relationship between CU&C and EFs has not been systematically investigated so far except in the case of Theory of Mind (ToM) tasks. In ToM tasks children must understand that others may hold a belief different from what they know is true. Young children's failure in false beliefs tasks has been explained on the grounds that young children lack the EF of inhibitory control (Hughes, 1998). Recently Zaitchik, Igbal and Carey (2013) examined relations between EF abilities and conceptual change in the domain of biology. They tested the hypothesis that the variance in the age at which children construct and deploy their first explicit theory of biology is due in part to individual differences in their EF abilities. Their results showed that in a group of 79 children ranging in age from 5 to 7 years, individuals' EF scores significantly predicted their scores in a battery of tasks assessing knowledge in biology, even after controlling for age and verbal IQ.

The relationship between EFs and conceptual change processes required in learning many science and mathematics concepts is interesting for a number of reasons. First, it would help to unify domain specific approaches to development and learning that focus on conceptual understanding and change and domain general approaches that focus on general mechanisms. EFs are domain general mechanisms that may be implicated in important ways in conceptual change processes. Second, the implication of EFs may explain why conceptual change learning is difficult, why they are substantial individual differences in the acquisition of science and mathematics concepts that require conceptual change, and why conceptual change is absent in some adult populations such as individuals with Williams Syndrome and patients with Alzheimer's who also have with impaired EFs (Zaitchik et al., 2013).

The purpose of the present research was to investigate the relation that EFs have with conceptual understanding and conceptual change in elementary school students who had some exposure to systematic science and mathematics instruction. We hypothesized that EFs, and particularly the EFs of inhibition and shifting, play an important role in the construction and re-organization of knowledge and thus in the acquisition of new concepts required in learning science and mathematics. We also hypothesized that low EFs may explain in part the significant individual differences observed in students' mathematics and science learning progressions and their abilities to profit from instruction.

In order to investigate these hypotheses we designed two computer-based, reaction time conceptual understanding and change (CU&C) tasks, which investigate the conceptual understanding and conceptual changes that take place with development and the learning of science and mathematics, and which were administered to 4^{th} and 6^{th} grade students: The *Re-Categorization (ReCat)* task and the *Sentence-Picture Verification Task (SPV)* task. The ReCat task investigates children's abilities to categorize words/concepts in initial and scientific categories. The SPV task assesses individuals' abilities to assess the truth or falsity of common-sense and scientific statements. Two additional computer-based, reaction time tasks were designed to test the EFs of Inhibition and Shifting. We predicted high correlations between performance in the EF and the CU&C tasks even when age and intellectual ability was partialed out. We also predicted that the children with low EFs would have particular difficulty with CU&C tasks and particularly with those requiring conceptual changes.

Method

Participants

The participants were 69 (29 boys, 40 girls) elementary school students, 33 4th graders ranging in age from 8.7 to 10.3 years (mean age: 9.7), and 36 6th graders ranging in age from 11.3 to 13.7 years (mean age: 11.8).

Tasks

The CU&C test battery consisted of two tasks: In *the Re-Categorization (ReCat) Task* a concept/word appeared on the computer screen and the children were asked to decide to which of two categories it is a member of. Three representative members were used to denote each of the two categories. Each word/concept appeared twice during the task, in two different categorization conditions: In the *Initial condition* the participants had to decide between an initial category and an anomalous category. In the *Scientific condition* the participants had to decide between an initial category and a scientific category. For example, they had to decide whether the concept *'force'* belongs with *'bicycle, motorcycle, car'* or *'strength, movement, push'* in the initial condition or with *'gravity, friction, reaction'* or *'strength, movement, push'* in the scientific condition. In both conditions, participants had to press the corresponding keyboard key for whichever category they thought that the target word was a member of. The items were counterbalanced for left/right responses. There were 4 practice trials with feedback. The categorization task included 28 target concepts, which appeared twice, so there were a total of 56 trials. The concepts were representative from 5 subject domains: Physics, Chemistry, Astronomy, Biology, and Mathematics.

In *The Sentence-Picture Verification (SPV) Task* the participants verified a total of 136 statements and pictorial representations. These stimuli provided 4 different explanations of 34 phenomena from five subject matter areas: Astronomy, Chemistry, Physics, Biology and Mathematics. Of the 4 different explanations, 2 were *consistent* with both initial and scientific views and 2 were *inconsistent* with one of them. For example, the participants were shown variations of the picture of a man and a spring (pulling, not pulling, stretching a lot or just touching and pulling the spring very little) and were asked to verify the sentences: *The man exerts force when the spring is pulled (Consistent, True), The man exerts force even when the spring is not pulled (Consistent, False) The man exerts force only when the spring is stretched (Inconsistent, False), The man exerts force even when the spring is not stretched (Inconsistent, True).* The participants were instructed to pay attention both to the verbal statement and to its pictorial representation. All initial and scientific alternatives were determined on the basis of existing research.

The EF task battery consisted of two Stroop-like tasks designed to test Inhibition and Shifting with three different stimuli, namely words, numbers and shapes. The tasks were developed for the needs of the current study and were adapted and presented using the E-prime software for computer administration. When the stimuli were words, the participants were instructed *to read the word aloud and press one of four buttons indicating the color in which the word was printed*. In the consistent condition (Con), the color word and the color in which it was printed were the same (the word RED was printed in red) while in the incongruent condition (Incon) the color word and the color in which it was printed were different (the word RED was printed in blue). In the case of numbers, the stimuli represented one of 4 numbers (1, 2, 3, or 4) presented in 1 to 4 repeated digits (for example, 2 could appear as 2, 22, 222, or 2222). The participants were instructed *to read the number presented on the screen aloud but press one of four buttons indicating the number of digits*. In the case of shapes, the stimuli presented a large shape consisting of smaller shapes. For example a triangle could consist of small triangles (consistent) or small circles (inconsistent). There were a total of 4 shapes used (triangle, circle, square and diamond). The participants were asked to *name the large shape but to press the response button indicating the smaller shape*. Each inhibition tasks included a total 90 trials - 45 for the congruent and 45 for the incongruent condition equally distributed between the three symbolic systems.

The same stimuli where used in the Shifting task as in the Inhibition task with the exception that in the Inhibition task the participants had to follow one specific rule while in the case of Shifting they had to shift between two rules. For example, when the stimuli were words, the participants had to shift between pressing the button to indicate the color of the word vs. the color in which the word was printed, depending on the instructions appearing on the screen of the computer. In the case of numbers they had to shift between naming the number or the number of digits in which the number was presented, and in the case of shapes they had to shift between naming the big geometrical shape or the small geometrical shapes from which the big geometrical shape was composed. Letters at the bottom of the screen indicated which rule was to be used. In the Con condition, participants had to respond to two consequent trials on the basis of the same rule, while in the Incon condition participants had to respond to two consequent trials on the basis of the different rule. The shifting tasks included a total of 96 trials, 48 for each condition. To assess intelligence a non-verbal IQ test, the Standard Progressive Matrices (SPM, Raven, Court & Raven, 1985), was used.

Procedure

Testing took place in the school's computer lab in a total of 5 sessions, 30 minutes each. All tasks were presented using the E-Prime v2.0 software (Schneider et al., 2012) in a single block, in a pre-randomized order.

Results

In all the analyses the dependent measures used were the mean percent accurate responses (ACC) and the mean RTs, in the initial and scientific trials of the CU&C task, and the consistent and inconsistent trials of the EF tasks. Responses with reaction times slower than 400ms in the EF tasks and slower than1000ms for the CU&C tasks were eliminated from the dataset. Because the two CU&C tasks (ReCat and SVP) were highly correlated both in accuracy and RTs we combined their scores into four aggregate CU&C measures: CU&C initial ACC (α=.88), CU&C initial RT (α=.89), CU&C scientific ACC (α=.46), and CU&C scientific RT (α=.75).

Four separate Principle Component Analyses with varimax rotation were run on the inconsistent and consistent responses on the EF tests for ACC and RT. All the analyses revealed one factor explaining (a) 55.87% of the total variance in the case of the EF inconsistent accuracy score (EFincon ACC, α = .83), (b) 42.00% of the total variance (α=.71) in the case of the EF inconsistent RT (EFincon RT), (c) 55.67% of the total variance (α=.811) in the case of the EF consistent accuracy (EFcon ACC), and (d) 41.46% of the total variance (α=.57) in the case of the EF consistent RT (EFcon RT). This allowed us to create four aggregate scores representing accuracy where the higher ACC scores indicate higher EFs, and where the lower RT scores indicate higher EFs. See Table 1for the means and SD.

Table 1: Mean percent accurate response and mean RTs for the aggregated CU&C and EF tasks.

Task	CU&C			
	Accuracy		Reaction times	
	M	SD	M	SD
Initial	80.72	13.18	4400.16	1188.74
Scientific	46.03	7.91	5276.98	1471.52
Task	**EF**			
	Accuracy		Reaction times	
	M	SD	M	SD
Consistent	75.48	14.47	1884.98	419.39
Inconsistent	42.92	7.80	2067.81	497.95

Significant correlations were obtained between the aggregate CU&C and EF ACC and RT scores as predicted. The correlations remained significant even when the Intellectual Ability (IA) score, as this has been measured with the RAVEN matrices and Grade were partialed out (Table2).

Table 2: Correlations between the CU&C and EF tasks when controlling for Grade and IA on ACC and on RT.

Accuracy (ACC)

	CU&Cin	CU&Csc	EFcon	EFincon
CU&Cin	1			
CU&Csc	.382**	1		
EFcon	.542**	.360**	1	
EFincon	.539**	.348**	.876**	1

**Correlation is significant at 0.01 level

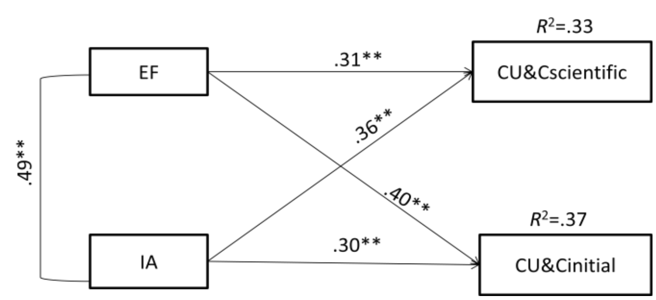
Reaction Times (RT)

	CU&Cin	CU&Csc	EFcon	EFincon
CU&Cin	1			
CU&Csc	.873**	1		
EFcon	.523**	.534**	1	
EFincon	.551**	.605**	851**	1

**Correlation is significant at 0.01 level

Two path-analytic models were applied to test the dynamic relations between CU&C, EF, and IA (RAVEN's score) using the structural equation modeling program EQSWIN (Bentler, 2003). In the first model, the EFincon ACC score was used as a measure of EF (i.e., the percentage of accurate scores on the inconsistent trials of the aggregate EF on the inhibition and shifting tasks). The target model specified direct relations of EFs and IA to CU&C initial ACC, and CU&C scientific ACC, while EF and IA allowed being correlated each other (see Figure 1). That is, the two types of conceptual understanding (demanding or not conceptual change) are regressed on the EFs as well as on IA, while the correlation between EFs and IA indicate their sharing common variance. The model demonstrated an acceptable fit, χ^2 (1, N=69) = 9.018, p = .003, CFI = .901, RMSEA = .343, (90% CI = .163 - .560). Figure 1 shows the resulting path coefficients of the model for our participants and the correlations between the EFs and IA. IA was positively correlated with EFs (r=.49, p < .001). CU&C scientific ACC was explained by IA (β = .32, p < .001) and EF (β = .31; p< .001; R^2 = .33). Additionally, CU&C initial ACC was explained by IA (β = .30, p < .001) and EF (β = .40; p< .001; R^2 = .37).

Figure 1: Executive Functions and Intelligence Ability as Predictors of CU&C.



Note: ** p < .001.

In the second model the EFincon RT was used as a measure of EF, and the CU&C initial and CU&C scientific ACC scores as a measure of CU&C. This model had a weak fit on the data, χ^2 (1, N=69) = 13.369, p < .001, CFI = .757, RMSEA = .426, (90% CI = .242 - .638).

To examine the second hypothesis, namely whether the level of EFincon ACC could predict the level of CU&C initial and CU&C scientific ACC, a prediction analysis (Froman & Hubert, 1980) was applied on the data. Participants at the higher percentile of EFincon ACC or of CU&C scientific and CU&C initial ACC respectively were classified as high achievers in the given variable, those who were found on the lower percentile as low achievers, and the rest as medium achievers. We assumed high or medium EFincon ACC is necessary to achieve a similar level of conceptual change (see Table 3 below). The predictions tested are indicated in the first three columns of Table 3

(Predicted level of CU&C accuracy). Based on Froman and Hubert (1980), cells marked with 0 indicate that they are consistent with the predictions while cells marked with 1 are not consistent with the predictions. The mark of 0.5 indicates that it is acceptable to find some participants in these cells. The observed distribution of participants is indicated in the remaining columns.

Table 3: Bivariate Frequency Distribution between EFs and CU&C

EF ACC Level	Predicted level on CU&C ACC			Observed CU&C scientific ACC			Observed CU&C initial ACC		
	L	M	H	L	M	H	L	M	H
L	0	1	1	15	4	4	15	5	3
M	0.5	0	1	5	13	5	7	8	9
H	1	0.5	0	2	7	14	2	9	12
Total				23	23	23	23	23	23

The results showed that the level of EFincon ACC could predict 44% better the achievement on the CU&C scientific ACC than if we assumed independence (▽=. 44, U = .608, z = 6.625, z > 1.96, p >.005). EFincon ACC could predict only 29.2% better the achievement on the CU&C initial ACC (▽=. 292, U = .615, z = 8.22, z > 1.96, p >.005) than if we assumed independence. EFincon ACC predicted better the CU&C scientific ACC than the CU&C initial ACC performance, but the difference was not statistically significant (z = 1.14, z < 1.96).

A similar prediction analysis using the EFincon RT as the level of EF did not produce a significant prediction on CU&C initial or CU&C scientific ACC. On the other hand, the analyses showed that the level of IA (scores on the RAVEN) could predict 29.7% better the achievement on the CU&C scientific than if we assumed independence (▽=. 297, U = .604, z = 4.58, z > 1.96, p >.005), and 32% better the achievement on the CU&C initial (▽=. 320, U = .607, z = 8.00, z > 1.96, p >.005) than if we assumed independence. Although IA predicted better the CU&C initial than the CU&C scientific ACC performance, this difference was not statistically significant (z = .219, z <1.96).

Discussion

The results demonstrated that performance in the EF and CU&C tasks is highly correlated, thus supporting the hypothesis that EFs are implicated in the conceptual understanding and conceptual change processes involved in science and mathematics learning. The correlations obtained were between the accuracy scores of the aggregate CU&C and EF tasks as well as between their RTs.

Accuracy scores on the EFs were interpreted to indicate that the participants could execute the inhibition and shifting tasks, and RTs to indicate how fast or easily they could do so. In many executive function studies, and particularly those used with adult participants, RTs are used as a measure of EFs because the tasks are very easy and there is little difference in accuracy between the consistent and inconsistent conditions The EF tasks that we employed, however, although relatively simple, were not very easy for the children. There were significant differences in accuracy between the consistent and the inconsistent aggregate EF scores (Table 1) that allowed us to use the accuracy scores as a good measure of EF (see also Davidson et al., 2006 for similar results with children).

The CU&C tasks allowed us to measure not only accuracy but also RTs. In all the studies that we know of that investigate the relationship between academic performance or conceptual change and EFs (e.g., van de Sluis, de Jong, & van der Leij, 2007; Zaitchik, et al., 2013) only accuracy scores are used and compared to EF accuracy or speed. In contrast, in the present study it was possible to directly compare the CU&C accuracy scores with the EF accuracy scores and the CU&C RTs with the EF RTs.

Accuracy scores on the CU&C tasks were interpreted to indicate the level of conceptual understanding and conceptual change the participants had achieved. More specifically, performance on the CU&C initial tasks was interpreted to indicate initial conceptual understanding, while performance on the CU&C scientific tasks to indicate the achievement of conceptual change. We interpreted the reaction time scores (RTs on accurate responses only), to indicate how fast the participants can recruit or express their initial or scientific understanding.

The results of the correlation analysis showed that there were high correlations between performance in the CU&C and EF tasks both in accuracy as well as in RTs. Indeed, the CU&C tasks themselves seemed to function like Executive Function tasks. In other words, performance in the tasks consistent with initial understanding was more accurate and faster compared to performance in the tasks requiring conceptual change and scientific understanding (Table 1). These finding are consistent with the existing literature (Shtulman & Valcarcel, 2012; DeWolf & Vosniadou, 2014) and support the hypothesis that initial conceptions are not replaced by scientific ones but continue to exist and to inhibit access to scientific responses.

There were only low or non-significant correlations between the EFincon RT and CU&C accuracy. This finding was clarified better from the path models applied on the data. The path-analytic model that explained our data was the one that used the aggregated EFincon ACC score as a predictor of CU&C accuracy rather than the model where the aggregated EFincon RT score was used as an index of EF efficiency. Methodologically these findings indicate that the EF tasks were relatively difficult for the children and therefore accuracy was the best measure of EFs instead of RT. At the theoretical level, they probably show that EFincon ACC and EFincon RT should be treated as two distinct aspects of EF performance with the EFincon ACC score serving as the best indicator of EF across young participants, (see also Davidson & et al., 2006). No matter the case our data supply clear evidence for the crucial role of EF in conceptual change but also in conceptual understanding.

The results also supported the hypothesis that children with low EFs would have difficulty with the CU&C tasks, particularly those requiring conceptual change. The prediction analysis showed that children who are in the low EF group are much more likely to be low CU&C achievers than high or even medium CU&C achievers. The measures of EF were better predictors of performance in the CU&C scientific tasks, whereas the measures of IA were better predictors of performance in the CU&C initial tasks. This finding suggests the greater involvement of EF compared to general IA in conceptual change processes, and point to the importance of EF in tasks that require abstraction, conceptual organization and particularly conceptual change. They are also important in explaining the significant individual differences we find in children's abilities to profit from instruction when conceptual changes are required.

This is a preliminary study and the results need to be replicated with a larger sample. Although the path analysis was justified on the grounds that the number of participants was more than five times the number of parameters analyzed in the path model, the overall sample size does not meet the requirement for a minimum of 100 participants (Kline, 1998). Further research is also needed to investigate developmental changes in performance in the CU&C tasks with a sample involving different ages in order to examine whether EF RTs become more predictive of CU&C accuracy with older participants who would be expected to find the EF tasks much easier. It would also be interesting to study the tradeoff between accuracy and RT with development. Preliminary results from the investigation of the developmental changes in our sample show that the 6th grade children were more accurate than the 4th grade children in both the initial and scientific CU&C responses, but there were no developmental effects on the RTs. Finally, it is important to investigate in further studies with older participants whether CU&C scientific responses become as accurate and fast as initial ones with the acquisition of expertise.

Acknowledgments

This research has been co-financed by the European Union (European Social Fund – ESF) and Greek national funds through the Operational Program "Education and Lifelong Learning" of the National Strategic Reference Framework (NSRF) - Research Funding Program: THALIS –UOA.

References

Bentler, P. M. (2003). EQS Structural Equations Program for Windows (Version 6.1 Beta Build 70) [Computer software]. Encino, CA: Multivariate Software.

Blair, C., & Razza, R. P. (2007). Relating effortful control, executive function, and false belief understanding to emerging math and literacy ability in kindergarten. *Child development*, *78*(2), 647-663.

Carey, S. (2009). *The origin of concepts*. New York: Oxford University Press.

Chi, M. T. H. (2008). Three types of conceptual change: belief revision, mental model transformation, and categorical shift. In S. Vosniadou (Ed.), *Handbook of research on conceptual change* (pp. 61-82). Hillsdale, NJ: Erlbaum.

Davidson, M. C., Amso, D., Anderson, L. C., & Diamond, A. (2006). Development of cognitive control and executive functions from 4 to 13 years: Evidence from manipulations of memory, inhibition, and task switching. *Neuropsychologia*, *44*(11), 2037-2078.

DeWolf, M., & Vosniadou, S. (2015). The representation of fraction magnitudes and the whole number bias reconsidered. *Learning and Instruction*, *37*, 39-49

Dunbar, K., Fugelsang, J., & Stein, C. (2007). Do naive theories ever go away? Using brain and behavior to understand changes in concepts. In M. Lovett & P. Shah (Eds.), *Thinking with data* (pp. 193–206). New York: Lawrence Erlbaum Associates.

Froman, T., & Hubert, L.J. (1980). Application of prediction analysis to developmental priority. *Psychological Bulletin*, 136-146.

Hughes R. (1998). Considering the vignette technique and its application to a study of drug injecting and HIV risk and safer behaviour. *Sociology of Health and Illness 20*, 381–400.

Kline, R. B. (1998). *Principles and practice of structural equation modeling*. NY: Guilford Press.

Miyake, A., Friedman, N. P., Emerson, M. J., Witzki, A. H., *&* Howerter, A. *(*2000*).* The unity and diversity of executive functions and their contribution to complex "frontal lobe" tasks: A latent variable analysis. *Cognitive Psychology*, *41*, 49–100.

Ni, Y., & Zhou, Y-D. (2005). Teaching and learning fraction and rational numbers: The origins and implications of whole number bias. *Educational Psychologist*, *40*(1), 27–52.

Schneider, W., Eschman, A., and Zuccolotto, A. (2012). E-Prime User's Guide. Pittsburgh: Psychology Software Tools, Inc.

Shtulman, A. & Valcarcel, J. (2012) Scientific knowledge suppresses but does not supplant earlier intuitions. *Cognition*, 124, 209-215.

Stafylidou, S., & Vosniadou, S., (2004). The Development of Students' Understanding of the Numerical Value of Fractions. In L. Verschaffel and S. Vosniadou (Guest Editors), Conceptual Change in Mathematics Learning and Teaching, Special Issue of *Learning and Instruction. 14*(5), pp. 503-518.

Van de Sluis, S., De Jong, P.F., & Van der Leij, A. (2007). Executive functioning in children, and its relations with reasoning, reading, and arithmetic. *Intelligence*, *35* pp. 427–449.

Vosniadou, S., & Verschaffel, L., (2004). Extending the Conceptual Change Approach to Mathematics Learning and Teaching. In L., Verschaffel and S. Vosniadou (Guest Editors), Conceptual Change in Mathematics Learning and Teaching, *Special Issue of Learning and Instruction*, *14*(5), pp. 445-451.

Zaitchik, D., Iqbal, Y., & Carey, S. (2013). The effect of executive function on biological reasoning in young children: An individual differences study. *Child Development*, 00, 1–16.