UC Merced
Proceedings of the Annual Meeting of the Cognitive Science Society

Title
Training flexible categorization to improve arithmetic problem solving: A school-based intervention with 5th graders

Permalink
https://escholarship.org/uc/item/1sw734r9

Journal
Proceedings of the Annual Meeting of the Cognitive Science Society, 44(44)

Authors
Iacono, Eric
Gros, Hippolyte
Clément, Evelyne

Publication Date
2022

Copyright Information
This work is made available under the terms of a Creative Commons Attribution License, available at https://creativecommons.org/licenses/by/4.0/

Peer reviewed

eScholarship.org Powered by the California Digital Library
University of California

Training flexible categorization to improve arithmetic problem solving: A school-based intervention with 5th graders

Eric Iacono (eric.iacono@cyu.fr)
Paragraphe Lab,
CY Cergy Paris University & Paris 8 University

Hippolyte Gros (hippolyte.gros@cyu.fr)
Paragraphe Lab,
CY Cergy Paris University & Paris 8 University

Evelyne Clément (evelyne.clement@cyu.fr)
Paragraphe Lab,
CY Cergy Paris University & Paris 8 University

Abstract

Because of its importance in academic achievement, especially in mathematics, training cognitive flexibility at school is a major issue. The present research investigates the effectiveness of a school-based intervention to improve proportion arithmetic problem solving. The study was conducted with 5th graders of 10 classes from 5 high-priority education schools in the Paris region. Students of the control and experimental groups took part in 8 learning sessions about proportion problem solving. The experimental group's training focused on comparing and flexibly categorizing the problems in the hopes to help students achieve a deeper understanding of proportion problems. Results show that training flexible categorization allowed the experimental group to progress more than the control group, in both categorization and solving tasks. The educational implications of our results are discussed.

Keywords: cognitive flexibility; school-based intervention; categorization; evidence-based education; arithmetic word problems

Introduction

Can a school intervention based on flexible categorization successfully target the expression of cognitive flexibility in arithmetic problem solving? By making it possible to adapt to a constantly changing environment, to discover creative solutions, to transfer knowledge from one situation to another, to infer the meanings of new words, or more generally to switch from one behavior to another depending on the environmental constraints, cognitive flexibility is considered a hallmark of human cognition (*e.g.*, Deák, 2003; Ionescu, 2012, 2017). While there is no unified definition in the literature, it is widely accepted that flexibility plays a key role in the development of thinking, language, reasoning, and knowledge acquisition (*e.g.*, Blaye & Bonthoux, 2001; Clément, 2009, 2022; Deák, 2003; Sloutsky & Fisher, 2008). Cognitive flexibility is often conceived of as the ability to cleverly change behavior in an appropriate manner depending on what a situation requires. However, upon closer scrutiny, this general definition of flexibility appears almost indistinguishable from a broad definition of intelligence. To overcome this limitation, cognitive scientists have striven to describe the specific cognitive processes underlying the flexible expression of our behavior and thoughts in terms of cognitive processes.

In a canonical sense, cognitive flexibility (also called set shifting, see Diamond, 2016) is described in the literature as an executive function consisting in the ability to shift attention from one stimulus, rule, or task to another (*e.g.*, Diamond, 2013; Lehto et al., 2003; Miyake et al., 2000; Meiran, Chorev, & Sapir, 2000; Monsell, 2003). However, it appears that in complex activities such as problem solving, cognitive flexibility also entails the ability to flexibly consider the same problem from different perspectives, to flexibly categorize it at different levels of abstraction. In the following sections, we present the empirical studies that investigate these two fundamental processes (*i.e.*, attentional shifting and flexible categorization).

Attentional shifting

Since the pioneer work of Miyake et al. (2000), cognitive flexibility, inhibition, and working memory are considered as the three core functions of executive control. In this approach, flexibility is often described as a *set shifting* (for a developmental approach, see for example Davidson et al., 2006). In that sense, the cognitive processes underlying set shifting are mainly related to attentional flexibility (Ionescu, 2017). In the following, we will refer to this process as attentional flexibility.

In a large body of research, attentional flexibility has been operationalized within the well-known task-switching paradigm (see Monsell, 2003). A series of tasks have been designed according to this paradigm, among which the widely used rule-based switching tasks. These consist in presenting multidimensional stimuli that first need to be categorized according to a specific dimension (*e.g.*, size) before changing the rule and switching to another dimension (*e.g.*, number of items). The Wisconsin test (Heaton et al.,

In J. Culbertson, A. Perfors, H. Rabagliati & V. Ramenzoni (Eds.), *Proceedings of the 44th Annual Conference of the Cognitive Science Society*. ©2022 The Author(s). This work is licensed under a Creative Commons Attribution 4.0 International License (CC BY).

1993) for adolescents and adults, or the Dimensional Change Card Sort (DCCS) for preschoolers (Frye, Zelazo, & Palfai, 1995; Zelazo, 2006**)** are some of the most typical examples of rule-based switching tasks. For instance, in the standard DCCS, children have to sort bi-dimensional cards (fish and boats that are either red or blue) according to a first rule (*e.g.*, color) and then, according to a second rule (*e.g.*, shape).

Developmental research reports two robust results regarding attentional shifting. First, factor analyses with children show that attentional shifting emerges later than inhibition and working memory (*e.g.*, Carroll, Blakey, & FitzGibbon, 2016). Second, before 3 years old, children present a perseverative behavior when asked to perform rule-based switching tasks: after sorting the stimulus according to one dimension (*e.g.*, shape), they experience considerable difficulty sorting it according to another dimension (*e.g.*, color) (*e.g.*, Deák & Wiseheart, 2015; Holt & Deák, 2015; Legare et al., 2018; Zelazo, 2006).

Although the task switching paradigm is well suited to assess attentional flexibility, its ability to assess other components of flexibility is limited for at least two reasons. First, in several situations involving flexibility, such as, for instance, problem solving, there is no imposed rule change and one needs to spontaneously adopt a new perspective without any outside prompt. Second, most of the time, achieving the goal requires to flexibly re-interpret (re-represent) the situation, which goes beyond a mere switch of attention. Thus, representational flexibility relies both on the attentional processes required to take into account environmental cues (for instance, noticing when the strategy used is not optimal), and on the ability to flexibly categorize the problem by re-interpreting the situation it depicts.

Flexible categorization in problem solving

The idea that a solver's ability to adequately categorize a problem is related to their ability to find its solution is perfectly exemplified by the seminal work of Chi, Feltovich, and Glaser (1981) on the nature of expertise. In their study, they asked both novices and experts to sort a set of physics problems into as many categories as they liked. Their results showed that while novices tend to categorize problems based on their surface features (*e.g.*, sorting together all the problems mentioning a spring), experts base their categories on deeper abstract principles (*e.g.*, putting together problems that can be solved by applying the Law of Conservation of Energy). The process by which one may ignore the surface features of a problem and flexibly adopt a new point of view to find its solution is a notoriously difficult endeavor.

In fact, an increasing amount of literature has shown that disregarding the initial interpretation of a problem and flexibly categorizing it according to a more abstract perspective is a key step in solving counter-intuitive problems (Brissiaud & Sander, 2010; Ragni, Kola, & Johnson-Laird, 2018). Notably, interferences between one's experience and knowledge about the world and one's reasoning have been shown to greatly impact solving performance (*e.g.*, Bassok, Wu, & Olseth, 1995; Bassok, Chase, & Martin, 1998; Cosmides & Tooby, 1992; Gros, Thibaut, & Sander, 2021). For instance, it has been demonstrated that differences of difficulty between isomorphic problems (problems sharing the same solution) may be explained by the influence of prior knowledge on the way the situation is interpreted. When prior knowledge leads participants to initially think that an action is not possible, they will take a longer time to get around this self-imposed limit and they will make more mistakes while doing so (Clément & Richard, 1997). In a similar fashion, getting rid of the common misconception that a subtraction necessarily describes the removal of a part from the whole is an arduous task (Brissiaud & Sander, 2010). Adopting the more accurate conception according to which a subtraction describes the distance between two numerical values requires a form of conceptual flexibility relying on flexible categorization of the very concept of subtraction (Hofstadter & Sander, 2013).

In sum, in problem solving, it has been demonstrated that semantic recoding and re-categorization of the situation are the core cognitive processes underlying flexible responses (*e.g.*, Clément, 2009, 2022; Clerc & Clément, 2016; Gamo, Sander, & Richard, 2010; Gros & Gvozdic, 2022; Gros, Thibaut, & Sander, 2020; Scheibling-Sève, Pasquinelli, & Sander, 2020; Scheibling-Sève, Sander, & Pasquinelli, 2017). In fact, finding the solution to a problem usually means looking at the problem from different perspectives. By adopting a new, usually more abstract point of view on the situation, one may be able to identify the deep structure of the problem and, in doing so, find a previously hidden path to the solution (Clément, 2009; Gros et al., 2020).

Flexibility and academic achievement

Whether it be scientific research in education, psychology, or didactics, a growing body of evidence points towards executive functions (*i.e.*, inhibition, working memory, and set shifting) as crucial contributors to school achievements (*e.g.*, Agostino, Johnson, & Pascual-Leone, 2010; Monette, Bigras, & Guay, 2011; St Clair-Thompson & Gathercole, 2006; Stad et al., 2018; Yeniad et al., 2013). For instance, after controlling for inhibition, working memory, planning, and fluid intelligence, Magalhães et al. (2020) found that cognitive flexibility accounts for a significant amount of variance in literacy and mathematics outcomes across Grades 2, 4, and 6.

More specifically, regarding math-achievement, in a recent study Hästö *et al.* (2019) demonstrated that cognitive flexibility predicts students' performance on finals, notably in mathematics. Cognitive flexibility has been argued to be a major contributor to math-performance, since, when one way of solving a mathematical problem isn't working, students need to switch between different strategies, and re-interpret the situation (Gros, Sander, & Thibaut, 2019; Gros, et al., 2020; Sander & Richard, 2005; Vicente, Orrantia, & Verschaffel, 2007).

Considered as determinant factors in school achievement, a body of research focused on interventions for improving executive functions (*e.g.*, Blakey & Carroll, 2015; Diamond

& Ling, 2016; Mennetrey & Angeard, 2018). Surprisingly, very few studies have been conducted on cognitive flexibility as the ability to consider the problems from different perspectives. In fact, most studies are often devoted to training working memory, inhibitory control, or attentional shifting, and to evaluate the effectiveness of such training in different areas of daily life (*e.g.*, Diamond, 2013; Diamond & Ling, 2016). To date, only a limited number of studies have targeted the stimulation of flexibility as it manifests itself in complex cognitive activities, such as problem solving. Therefore, the present study focuses on the representational flexibility involved in proportional math problem solving.

The present study

Because of the crucial role of flexibility in school achievement, fostering it in the classroom is a major educational question. The aim of the present study is to promote flexible categorization in mathematics with 5th graders, and more particularly in arithmetic proportion word problems. In order to encourage students to identify, beyond their superficial similarities or dissimilarities, the abstract and deep structure of problems, we conceived an educational intervention based on comparison between proportion arithmetic word problems. In fact, in the analogical transfer literature, it is well established that comparing two analogous problems (sharing the same solution path) leads to a more abstract representation of which category the problems belong to (Catrambone & Holyoak, 1989; George & Wiley, 2018; Holyoak, 2005; Loewenstein, Thompson, & Gentner, 1999). In the same way, fostering students to compare two strategies to find the solution to the same problem allows them to flexibly re-interpret the problem in such a way that they encode the problem in a higher degree of abstraction (Brissiaud, 1994; Gamo et al., 2010; Scheibling-Sève, Sander, & Pasquinelli, 2017).

In accordance with these works, we postulate that encouraging children to compare problems and to identify their superficial and deep similarities will encourage them to perceive the abstract mathematical structure underlying proportion arithmetic word problems. This should help promote flexibility in the way students perceive, categorize, and solve the problems.

Hypotheses

We therefore hypothesized that the trained group (experimental group) would present a higher progression than the control group, both on categorization and resolution tasks. More precisely, at the post-test, the trained group should more frequently spontaneously categorize the problems according to their structural similarities rather than to their superficial resemblance, compared to the control group (Hypothesis 1). Second, the experimental group should display better post-test solving performance than the control group (Hypothesis 2).

Method

Participants

The experiment was conducted with 10 classes of 5th graders belonging to five elementary schools located in high-priority education networks in the Paris region. A total of 147 5th graders took part in the study (mean age = 11.4 years; SD = 0.45). The experimental group included 77 students (40 girls). The control group included 70 students (39 girls). The 5 classes of each group were drawn from the same high-priority education network, thus indicating comparable socioeconomic status and overall diversity. Written consent to take part in the experiment was obtained from all the participants' parents. The experiment was conducted during regular school hours for both groups.

Design

To assess the effectiveness of our educational intervention, the experiment was designed in three phases; a pretest, followed by 8 learning sessions on multiplicative problems (the content of which differed between the experimental and control groups), and a post-test. The learning sessions were of identical duration and frequency between the two groups. The pre- and post-test were identical, the only difference consisting in a supplementary task at the end of the post-test, in which students were asked to create word problems corresponding to the four problem structures that had been taught (see Table 1).

In an effort to maximize the ecological validity of the study, the training sessions in both groups were conducted by the classes' regular teachers. Prior to the study, the teachers had been taught how to use the experiment's educational materials to teach their classes. The first author was in charge of training the teachers, due to his background as an educational advisor specialized in teacher training. The teachers from the experimental group were introduced to the online platform that they would have to use to conduct the learning sessions. They were also given instructions to mainly focus training on a better understanding of the quaternary problems as defined by Vergnaud (1983). It should be noted that the teachers were completely blind to the hypotheses of our study, they did not know whether they were in the control or in the experimental group, and they had to follow the set timeline for each session.

Materials

Pre- and post-tests

A series of 16 problems was constructed according to the four multiplicative structures defined by Vergnaud (1983) : multiplication (*e.g.*, *In 1 box there are 107 sweets. I have 253 identical boxes. How many sweets do I have in total ?*), first-type division (*e.g.*, *I have 1 box of sweets to share between 24 people. It contains 72 pieces. How many sweets will each person get?*), second-type division (*e.g.*, *In 1 box I have 107 identical sweets. I have 749 sweets in total. How many boxes can I fill?*), and direct proportion (*e.g.*, *I fill 2 boxes with 30 sweets. I have 48 boxes. How many sweets do I need to fill* *them?*). Each of these four multiplicative structures was presented in four different contexts describing either sweets, flowers, notebooks, or photo albums (see Table 1). The length of the problems was controlled, each problem statement shared the same number of sentences and presented the three numerical values in the same order. In the categorization task, each problem was printed on a card, following Chi et al.'s (1981) experimental design. The participants were instructed to classify the problems as they saw fit, creating as many categories as they wished. Then, in the problem-solving task, students were presented with an 8-page booklet displaying 8 of these problems (one per page). Students were instructed to try to solve the problems. Problem order was randomized between participants.

Table 1: Sample problem statements belonging to the 4 relevant structures and presented in 4 different contexts

Multiplication	First type division	Second type division	Direct proportion
In 1 box there are 107 sweets. I have 253 identical boxes. How many sweets do I have in total ?	I have 1 box of sweets to share between 24 people. It contains 72 pieces. How many sweets will each person get?	In 1 box I have 107 identical sweets. I have 749 sweets in total. How many boxes can I fill?	I fill 2 boxes with 30 sweets. I have 48 boxes. How many sweets do I need to fill them?
I buy 1 bouquet of roses for each of my 3 children. The bouquet costs 17 euros. How much should I pay?	I buy 1 bouquet of 120 roses. I pay €276. How much does each rose cost?	1 bouquet contains roses that cost €3 each. I pay €69. How many roses are there?	I buy 3 roses in a shop. It costs me €12. How much do I have to pay if I want 98 roses?
I have 1 photo album with 13 pages. I can store 18 photos per page. How many photos can I store?	I have 1 photo album. Across 21 identical pages, I have a total of 126 photos. How many photos are there on each page?	In 1 album, I put 14 photos per page. I have 392 photos. How many pages are required to store them?	I organize my photos in 3 albums. They can fit 147 photos in total. How many albums are required for 343 photos?
In my school, 1 pupil receives 11 notebooks. I have 228 pupils. How many notebooks will I hand out?	In my class 1 pupil receives 4 notebooks. I hand out a total of 96 notebooks. How many pupils do I have?	For 1 row of tables, I hand out 13 notebooks. I have 91 notebooks. How many rows can I complete?	For every 3 pupils, I hand out 24 notebooks. There are 336 pupils in the school. How many notebooks do I need in total?

Learning sessions

Learning sessions took place in 8 sessions over 4 weeks (two 55 minutes sessions each week). Teachers of the experimental group were given direct online access to each lesson so they could display the educational materials onto the interactive whiteboard. This made it possible to ensure that the teachers could each teach the exact same lesson in their own way. Furthermore, we had printed out a sheet for each teacher reminding them of how to conduct the session. This was done so that the experiment was as close as possible to real-life conditions; the type of training and instructions that teachers received was similar to the guidelines they regularly receive following educational reforms. In the control group, the learning sessions followed the regular curriculum, using the teachers' usual Math textbook (*e.g.*, Charnay et al., 2017). These sessions were built around the proportionality table and the "rule of three". In the experimental group, learning sessions were based on a general principle of comparison between analogous problems instead.

The first 6 sessions were focused on the comparison of problems belonging to two multiplicative structures (*i.e.*, first-type division/second-type division; multiplication/ direct proportion; multiplication/ first-type division; second-type division/direct proportion; multiplication/second-type division; first-type division/ direct proportion). In the last two lessons, problems belonging to the four multiplicative structures were compared. An equivalent number of problems sharing or not the same surface similarities was shown in each lesson. The aim of the experimental intervention was to get the children to ignore the surface similarities between the problems and identify their deep structure.

Results

Categorization task

As in Gros et al. (2021), we used similarity analysis to investigate the categories created by the participants. We coded the categories created by each participant at pre-test and post-test with a co-occurrence matrix describing how many times two problem statements were grouped together within the same category. We used these co-occurrence matrices to compute 2 proximity matrices for each group, describing the mean perceived similarity between each problem at pre-test and post-test. These matrices provide, for each pair of problem statements, a numerical estimation of the frequency at which the problems were sorted together.

We calculated the mean proximity score of the 24 pairs of problems sharing the same structure (*e.g.*, multiplication problem mentioning sweets with multiplication problem mentioning notebooks) for each group and each test. At pre-test, the mean proximity score for problems sharing the same structure in the control group (m = 0.554, SD = 0.013) was not significantly different from that of the experimental group (m = 0.555, SD = 0.017): t(23) = 0.25, p = .805, paired t-test. On the other hand, at post-test, the mean proximity score for structural pairings was significantly higher in the experimental group (m = 0.601, SD = 0.018) than in the control group (m = 0.536, SD = 0.012)), as hypothesized (t(23) = 17.19, p < .0001, paired t-test). Thus, the data collected from students' spontaneous categories supports Hypothesis 1: the experimental group was keener to create categories based on the problems' structure than the control group, which relied more heavily on surface similarities in this task.

Solving task

To evaluate the progress of the control and experimental groups, we looked at the increase in performance (number of correctly solved problems) between the pre-test and the post-test (see Fig. 1). We removed from the progress analysis the participants who were unable to participate either to the pre-test or to the post-test. The analysis was performed on the remaining 110 students. Results showed that participants in the experimental group went from an average performance score of 33.73% at the pre-test to an average score of 58.13% at the post-test (t(62) =, p < .0001, paired t-test). The participants in the control group also progressed significantly, going from an average performance of 33.24% at pre-test to an average of 44.41% at post-test (t(46) =, p < .01, paired t-test). Crucially, the mean progress between pre and post-test was significantly higher in the experimental group (m = 24.40%) than in the control group (m = 11.17%), t(108) =, p < .01, independent t-test.

These results support Hypothesis 2: while both groups achieved similar performance at pre-test, the experimental group benefited more from the 4-week training. In other words, students who followed the learning sessions based on comparison between analogous problems were more likely to find the solution to the proportion problems than students who followed the textbook-based training focusing on proportionality tables and "rule of three".

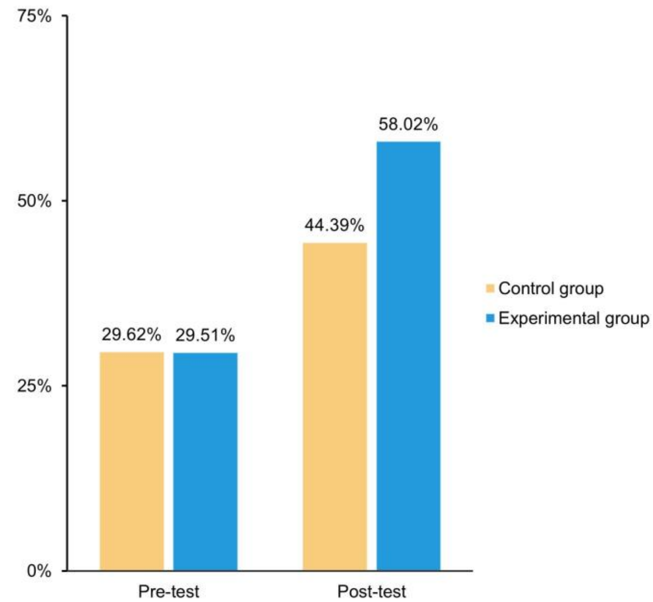

Fig 1. Students' performance in the solving task, before and after the intervention, depending on the training group.

Discussion

This study evaluated the effectiveness of an educational intervention targeting the understanding and solving of proportion problems. The intervention based on comparisons between analogous proportion problems sharing or not surface similarities, proved successful in improving performance both in problem categorization (Hypothesis 1) and in problem solving (Hypothesis 2). As predicted, there was no significant difference between the two groups' performance at pre-test, while students from the experimental group improved significantly more at post-test, both in their ability to categorize the problems based on their deep structure, and in their solving performance. In other words, participants in the experimental group developed a higher proficiency in proportion problem solving and they were able to flexibly categorize the problems using a representation with a higher degree of abstraction, moving away from the problems' superficial features.

These findings have important implications for pedagogical design. Indeed, while most French math textbooks teach about proportion problem solving by resorting to procedures (*e.g.*, how to use a proportionality table) and rules (*e.g.*, "the rule of three"), here the training focusing on problem comparison helped students achieve higher performances. Making comparisons helped them to see the structural commonalities between problems sharing the same solution principle and showed them which aspects of the problems were relevant and which could be disregarded. We believe that such an approach could be beneficial in other aspects of mathematical reasoning, as well as in other fields of knowledge. This comparison process makes it possible for students to learn to adopt a different point of view on the problems they encounter, flexibly categorizing the situation to identify the optimal way to broach and solve it. In this perspective, a sequence of comparisons between increasingly dissimilar problems sharing the same solution could result in a deeper understanding of the notions at hand, similarly to how *concreteness fading* (Fyfe, McNeil, Son, & Goldstone, 2014) uses examples of increasing abstraction to promote transfer.

Another promising direction for future research may lay in the study of the phenomenological components of the process by which students suddenly manage to flexibly recategorize a given problem. Indeed, recent developments in the insight literature have brought forth new venues to coin the subjective experience of learners going through "Aha!" moments (*e.g.*, Creswell et al., 2016; Laukkonen & Tangen, 2018; Webb, Little, & Cropper, 2016). "Aha!" experiences are thought to play a crucial role in helping students overcome their difficulties in learning mathematics (Liljedahl, 2007), and they are said to improve learning and memorization in general (Kizilirmak, Galvao Gomes da Silva, Imamoglu, & Richardson-Klavehn, 2016). Thus, identifying which training conditions lead to these feelings of insight may bring converging evidence regarding the relevance of similar school interventions based on flexible categorization.

Finally, this study highlights the importance of cognitive flexibility as a central lever of school learning, even outside of task-switching situations. As such, characterizing the different components of cognitive flexibility may be a key step in designing more effective school interventions. Indeed, while it may be argued that attentional flexibility has but a limited influence in class, the oft-forgotten representational flexibility appears to play a major part in the identification of problems' deep structure. This component of cognitive flexibility relying on flexible categorization seems to be a strong contender for school interventions aiming at improving near and far transfer.

Acknowledgements

This research was supported in part by a research grant (CYIn-AAP2021-AmbEm-0000000026) attributed by CY Initiative. We are grateful for the stimulating collaboration with schools Pierre Brossolette, Maurice Ravel, Jacqueline Auriol, Jules Ferry, and Paul Raoult, as well as for the helpful contributions of school teachers Frédérique Rouillé, Romain Grosjean, Rémy Boutrand, Sylvain Chebbance, Ivonig Thepot, Myriam Ibodé, Thomas Morel, Arnaud Morandi, François Alexandre Verger, Sandrew Calixto and all their students.

References

Agostino, A., Johnson, J., & Pascual-Leone, J. (2010). Executive functions underlying multiplicative reasoning: Problem type matters. *Journal of Experimental Child Psychology, 105(4)*, 286–305.

Bassok, M., Chase, V. M., & Martin, S. A. (1998). Adding apples and oranges: Alignment of semantic and formal knowledge. *Cognitive Psychology, 35*, 99–134.

Bassok, M., Wu, L. L., & Olseth, K.L. (1995). Judging a book by its cover: Interpretative effects of content on problem-solving transfer. *Memory & Cognition, 23*, 354–367.

Blakey, E., & Carroll, D. J. (2015). A short executive function training program improves preschoolers' working memory. *Frontiers in psychology*, 6, 1827.

Blaye, A., & Bonthoux, F. (2001). Thematic and taxonomic relations in preschoolers: The development of flexibility in categorization choices. *British Journal of Developmental Psychology*, 19(3), 395-412.

Brissiaud, R. (1994). Teaching and Development: Solving "Missing Addend" Problems Using Substraction. *European Journal of Psychology of Education, 9 (4)*, 343-365

Brissiaud, R., & Sander, E. (2010). Arithmetic word problem solving: a Situation Strategy First framework. *Developmental Science*, 13, 92-107.

Carroll, D. J., Blakey, E., & FitzGibbon, L. (2016). Cognitive flexibility in young children: Beyond perseveration. *Child Development Perspectives*, 10, 211–215.

Catrambone, R., & Holyoak, K. J. (1989). Overcoming contextual limitations on problem-solving transfer. Journal of Experimental Psychology: Learning, Memory, and Cognition, 15(6), 1147–1156.

Charnay, R., Combier, R., Oussuc, M. P., & Madier, O. (2017) *Cap Maths CM2, Manuel de l'élève*. Hatier.

Chi, M. T., Feltovich, P. J., & Glaser, R. (1981). Categorization and representation of physics problems by experts and novices. *Cognitive Science*, 5(2), 121-152.

Clément, E. (2009). *La résolution de problème. A la découverte de la flexibilité cognitive.* Paris: Armand Colin.

Clément, E. (2022). Successful solution discovery and cognitive flexibility. In E. Clément (Ed.), *Cognitive flexibility, the cornerstone of learning*, London: ISTE Edition.

Clément, E., & Richard, J.-F. (1997). Knowledge of domain effects in problem representation: The case of Tower of Hanoi isomorphs. *Thinking and reasoning*, 3, 133-157.

Clerc, J., & Clément, E. (2016). . Metacognition and cognitive flexibility in transfer of learning. In J. Benson (Ed.), *Metacognition: Theory, Performance and current.* New-York: Nova Science Publishers.

Cosmides, L., & Tooby, J. (1992). Cognitive adaptation for social exchange. In J.H. Barkow, L. Cosmides, & J. Tooby (Eds.), *The Adapted Mind: Evolutionary Psychology and the Generation of Culture.* Oxford: Oxford University Press.

Creswell, K., Layette, M., Schooler, J., Wright, A., & Pacilio, L. (2016). *Visceral states call for visceral measures: verbal overshadowing of hunger ratings across assessment modalities. Assessment 25*, 173–182.

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*, 2037-2078.
Deák, G. (2003). The development of cognitive flexibility and language abilities. *Advances in Child Development and Behavior*, *31*, 273-328.
Deák, G., & Wiseheart, M. (2015). Cognitive flexibility in young children: General or task-specific capacity? *Journal of Experimental Child Psychology, 138,* 31-53.
Diamond, A. (2013). Executive functions. *Annual Review of Psychology*, 64, 135–168.
Diamond, A. (2016). Why improving and assessing executive functions early in life is critical. In J. A. Griffin, P. McCardle, & L. S. Freund (Eds.), *Executive function in preschool-age children: Integrating measurement, neurodevelopment, and translational research. American Psychological Association.*
Diamond, A., & Ling, D. S. (2016). Conclusions about interventions, programs, and approaches for improving executive functions that appear justified and those that, despite much hype, do not. *Developmental Cognitive Neuroscience, 18*, 34-48.
Frye, D., Zelazo, D., & Palfai, T. (1995). Theory of mind and rule-based reasoning. *Cognitive Development, 10,* 483–527.
Fyfe, E. R., McNeil, N. M., Son, J. Y., & Goldstone, R. L. (2014). Concreteness fading in mathematics and science instruction: A systematic review. *Educational Psychology Review*, 26(1), 9-25.
Gamo, S., Sander, E., & Richard, J.-F. (2010). Transfer of strategies by semantic recoding in arithmetic problem solving. *Learning and Instruction, 20,* 400-410.
George, T., & Wiley, J. (2018). Breaking past the surface: Remote analogical transfer as creative insight. In Insight: On the Origins of New Ideas.
Gros, H., & Gvozdic, K. (2022). Context, content effects and flexibility. In E. Clément (Ed.), *Cognitive flexibility, the cornerstone of learning,* London: ISTE Edition.
Gros, H., Sander, E., & Thibaut, J.-P. (2019). When masters of abstraction run into a concrete wall: Experts failing arithmetic word problems. *Psychonomic Bulletin & Review*, 26(5), 1738-1746.
Gros, H., Thibaut, J.-P., & Sander, E. (2020). Semantic congruence in arithmetic: A new conceptual model for word problem solving. *Educational Psychologist*, 55, 1532–6985.
Gros, H., Thibaut, J.-P., & Sander, E. (2021). What we count dictates how we count: A tale of two encodings. *Cognition*, 212, 104665.
Hästö, P., Palkki, R., Tuomela, D., & Star, J. R. (2019). Relationship between mathematical flexibility and success in national examinations. *European Journal of Science and Mathematics Education, 7 (1)*, 1-13.
Heaton, R. K., Chelune, G. J., Talley, J.L., Kay, G., Curtiss, G., (1993). *Wisconsin Card Sorting Test Manual, revised* edition, Psychological Assessment Resources, Odessa, FL
Hofstadter, D., & Sander, E. (2013). *L'analogie. Cœur de la pensée*. Odile Jacob, Paris
Holt, A. E., & Deák, G. (2015). Children's task-switching efficiency: missing our cue?. *Journal of Cognition and Development*, 16(2), 261-285.
Holyoak, K. J. (2005). Analogy. In K.J. Holyoak & R.G. Morrison, (Eds.), *The Cambridge handbook of thinking and reasoning*. Cambridge: Cambridge University Press.
Ionescu, T. (2012). Exploring the nature of cognitive flexibility. *New Ideas in Psychology, 30,* 190-200.
Ionescu, T. (2017). The variability-stability-flexibility pattern: A possible key to understanding the flexibility of the human mind. *Review of General Psychology, 21(2),* 123–131.
Kizilirmak, J.M., Galvao Gomes da Silva, J., Imamoglu, F., & Richardson-Klavehn, A. (2016) Generation and the subjective feeling of "aha!" are independently related to learning from insight. *Psychological Research* 80, 1059–1074 (2016).
Laukkonen, R. E., & Tangen, J. M. (2018). How to detect insight moments in problem solving experiments. *Frontiers in psychology, 9*, 282.
Legare, C. H., Dale, M. T., Kim, S. Y., & Deák, G. (2018). Cultural variation in cognitive flexibility reveals diversity in the development of executive functions. Scientific Reports.
Lehto, J. E., Juujärvi, P., Kooistra, L., & Pulkkinen, L. (2003). Dimensions of executive functioning: Evidence from children. *British Journal of Developmental Psychology, 21*, 59–80.
Liljedahl, P. G. (2005) Mathematical discovery and affect: the effect of AHA! experiences on undergraduate mathematics students, *International Journal of Mathematical Education in Science and Technology, 36*:(2-3), 219-234.
Loewenstein, J., Thompson, L., & Gentner, D. (1999). Analogical encoding facilitates knowledge transfer in negotiation. *Psychonomic Bulletin & Review, 6,* 586-597.
Magalhães, S., Carneiro, L., Limpo, T., & Filipe, M. (2020). Executive functions predict literacy and mathematics achievements: The unique contribution of cognitive flexibility in grades 2, 4, and 6. *Child Neuropsychology*, 26(7), 934-952.
Meiran, N., Chorev, Z., & Sapir, A. (2000). Component processes in task switching. *Cognitive Psychology*, 41(3), 211-253.
Mennetrey, C., & Angeard, N. (2018). Cognitive flexibility training in three-year-old children. *Cognitive Development*, 48, 125-134.
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.
Monette, S., Bigras, M., & Guay, M.-C. (2011). The role of the executive functions in school achievement at the end of Grade 1. *Journal of Experimental Child Psychology, 109*, 158-173.

Monsell, S. (2003). Task switching. *Trends in Cognitive Sciences*, 7(3), 134–140.

Ragni, M., Kola, I., & Johnson-Laird, P. N. (2018). On selecting evidence to test hypotheses: A Theory of Selection Tasks. *Psychological Bulletin, 144* (8), 779-796.

Sander, E., & Richard, J.-F. (2005). Analogy and transfer: encoding the problem at the right level of abstraction. In *Proceedings of the 27th Annual Conference of the Cognitive Science Society*. Stresa, Italy (pp. 1925-1930).

Scheibling-Sève, C., Pasquinelli, E., & Sander, E. (2020). Assessing conceptual knowledge through solving arithmetic word problems. *Educational Studies in Mathematics*, 103, 293-311

Scheibling-Sève, C., Sander, E., & Pasquinelli, E. (2017). Developing cognitive flexibility in solving arithmetic word problems. In B. C. Love, K. McRae, & V. M. Sloutsky (Eds.), *Proceedings of the 39th Annual Meeting of the Cognitive Science Society* (439-444). Austin, TX.

Sloutsky, V. M., & Fisher, A. V. (2008). Attentional learning and flexible induction: How mundane mechanisms give rise to smart behaviors. *Child Development*, 79(3), 639-651.

Stad, F. E., Van Heijningen, C. J. M., Wiedl, K. H., & Resing, W. C. M. (2018). Prediction school achievement: Differential effects of dynamic testing measures and cognitive flexibility for math performance. *Learning and Individual Différences*, *67*, 117-125.

St Clair-Thompson, H. L., & Gathercole, S. E. (2006). Executive functions and achievements in school: Shifting, updating, inhibition, and working memory. *Quarterly Journal of Experimental Psychology, 59(4)*, 745–759.

Vergnaud, G. (1983). Multiplicative structures. In R. Lesh & M. Landau (Eds.), *Acquisition of Mathematics Concepts and Processes*. New York: Academie Press.

Vicente, S., Orrantia, J., & Verschaffel, L. (2007). Influence of situational and conceptual rewording on word problem solving. *The British journal of educational psychology*, *77*, 829–848.

Webb, M. E., Little, D. R., & Cropper, S. J. (2016). Insight is not in the problem: Investigating insight in problem solving across task types. *Frontiers in psychology*, *7*, 1424.

Yeniad, N., Malda, M., Mesman, J., van Ijzendoorn, M. H., & Pieper, S. (2013). Shifting ability predicts math and reading performance in children: A meta-analytical study. *Learning and Individual Differences, 23*, 1–9.

Zelazo, P. D. (2006). The Dimensional Change Card Sort (DCCS): a method of assessing executive function in children. *Nature Protocols*, 1(1), 297–301.