Planning, Execution, and Revision in Mathematics Problem Solving: Does the Order of the Phases Matter?

Trinidad García[a,b], Jan Boom[c], Evelyn H. Kroesbergen[c], José Carlos Núñez[a] and Celestino Rodríguez[a]*

[a]Department of Psychology, University of Oviedo, Oviedo, Spain; [b]Faculty Padre Ossó, Oviedo, Spain; [c]Department of Social and Behavioural Sciences, Utrecht University, Utrecht, The Netherlands

Trinidad García.

Address: Faculty of Psychology. Plaza Feijoo s/n, 33003 Oviedo (Asturias), Spain.

Jan Boom.

Address: Faculty of Social and Behavioural Sciences. Heidelberglaan 1, 3584 CS Utrecht, The Netherlands.

Evelyn H. Kroesbergen.

Address: Faculty of Social and Behavioural Sciences. Heidelberglaan 1, 3584 CS Utrecht, The Netherlands.

José Carlos Núñez.

Address: Faculty of Psychology. Plaza Feijoo s/n, 33003 Oviedo (Asturias), Spain.

*Celestino Rodríguez.
Address: Faculty of Psychology. Plaza Feijoo s/n, 33003 Oviedo (Asturias), Spain.
E-mail: rodriguezcelestino@uniovi.es

*Corresponding autor

Acknowledgments

This work was funded by the Council of Economy and Employment of the Princedom of Asturias (Spain) (Ref. FC-GRUPIN-IDI/2018/000199). The authors acknowledge the collaboration of the management teams and teachers of the educational centers participating in the study.

Planning, Execution, and Revision in Mathematics Problem Solving: Does the Order of the Phases Matter?

The present study analysed the mathematical problem-solving processes, in terms of linearity and recursion, and the relationship with actual and self-perceived performances of a sample of 524 students of upper-elementary students. The results showed a more linear than recursive process while performing the tasks, mainly characterized by continuity. The use of planning strategies before execution and the use of revision strategies after this phase were both significantly related to good performance, even if rates of success were low. The presence of a linear and hierarchical resolution process was related to students´ judgments of success, while recursion, or going back in the process, was associated with judgments of failure. Results are discussed in the light of current research on mathematics problem-solving.
Keywords: Mathematics Education; Problem Solving; Elementary School; Process; Self-perception

Introduction

Solving mathematical problems is a common task for students at all educational levels. Problem solving is a core goal of mathematics instruction at school, which is justified by the great importance that this skill has in everyday life and in the workplace. Although solving problems efficiently can be extremely beneficial in both contexts, previous research shows that students frequently struggle with these tasks (Babakhania, 2011; García, Rodríguez, González-Castro, González-Pienda, & Torrance, 2016; Silver, Ghousseini, Gosen, Charalambous, & Font Strawhun, 2005). There are different definitions of what constitutes “mathematical problem-solving”, most of which emphasize the complex nature of this activity. According to Raynal and Rieunier (1997; in Căprioarăa, 2015), problem-solving simultaneously mobilizes intellectual faculties such as memory, perception, reasoning, conceptualization, language, as well as emotional control, motivation, self-confidence and monitoring. On the other hand, Schoenfeld (1992) state that problem solving implies “thinking mathematically”, which involves mathematics core knowledge; problem-solving strategies such as monitoring and control; effective use of one´s resources; having a mathematical perspective; and engagement in mathematical practices (p. 335). In fact, successfully solving mathematical problems has been proven to rely on affective-motivational, cognitive, self-regulatory and metacognitive components constantly interacting with each other (Babakhania, 2011; Căprioarăa, 2015; García, Betts, González-Castro, González-Pienda, & Rodríguez, 2016; Jitendra, Dupuis, & Zaslofsky, 2014; Schoenfeld, 1992). More specifically, students who struggle with solving mathematical problems are poor at effectively implementing metacognitive and self-regulatory strategies. On the same lines, according to recent studies, there is also an important relationship between metacognitive and self-regulatory components and students' own judgments of performance (or self-perceived performance) (Dunlosky & Rawson 2012; Dunlosky & Thiede, 2013; Finn & Metcalfe 2014; García, Rodriguez et al., 2016; Hacker, Bol, & Bahbahani, 2008; Laua, Kitsantasb, & Millerc, 2015). These studies suggest that those students who are more accurate in their judgments commonly show higher levels of metacognitive control over their own learning processes.

The studies above provide initial evidence on the important link between the process, actual performance, and self-perceived performance in mathematics problem-solving; a link that must be more deeply examined. However, there is an important factor that must also be considered, which is how the process involved in solving mathematics problems is organized. From a Self-Regulated Learning based perspective (Zimmerman, 2000, 2008), learning is characterized by a process where planning, execution and revision phases alternate with each other in a cyclical –recursive- relationship (Carlson & Bloom, 2005). However, from the perspectives of other models, such as those by Polya (1981) and Mayer (2003), learning -especially problem-solving- is defined as being a linear (rather than a recursive) cognitive process; wherein planning, execution and revision seem to follow each other in a sequential order. Considering these two possibilities, and given the lack of empirical evidence on this issue to date, the current study examines the process involved in solving mathematical problems to see whether they can be better understood in terms of recursion or of linearity, and to analyze whether this process is related to students' actual and self-perceived performance in problem-solving.

Mathematical problem-solving as a process

Current problem-solving models focus on the process rather than on the content aspects of mathematics ~~(Schoenfeld, 1992).~~ According to Schoenfeld (1992), problem solving ~~is based on observation and~~ implies searching for patterns, which involves abstraction, symbolic representation, and symbolic manipulation as main tools. It also implies a cyclic process that goes from data to deduction to application and that is repeated each time we face a mathematical problem.

There are some previous studies that have examined the process involved in solving mathematical problems at different educational stages (García, Rodríguez et al., 2016; García, Betts et al., 2016; Jacobse & Harskamp, 2012; Tambychika, Subahan, & Meerahb, 2010; Verschaffel et al., 1999). These studies showed that difficulties in solving mathematical problems may occur at any phase during performance (i.e., planning-execution-evaluation; Zimmerman, 2000), with the phases of planning and evaluation commonly regarded as more problematic. In this sense, students commonly demonstrate difficulties in planning how to execute the problem-solving, using inadequate or insufficient strategies and devoting their efforts to performing calculations; while, for most of them, the third phase of evaluation seems unnecessary. Thus, most students do not usually complete the whole problem-solving process, which commonly leads to poor performance. In this same line, authors such as De Bock, Verschaffel and Janssens (1998), or De Corte and Somers (1982) found that when facing unfamiliar complex mathematical problems, students usually do not apply effective strategies such as organizing the information though a drawing or sketch, dividing the problem into different parts, or guessing and checking. Also, self-regulatory and metacognitive strategies are commonly scarce (i.e. analysing the problem, monitoring the solution process, evaluating its outcome, etc.). Instead, students frequently jump to ~~do~~ calculations, without considering any other alternatives even when performing mathematical operations may not be working. Analysing this sometimes erratic process is important, because a crucial relationship between 'process and product' (i.e. resulting performance on the task) has been extensively reported (Tambychika et al., 2010).

In this same line, Verschaffel et al. (1999) state that, despite the formal education and training provided by school-teachers in the area of mathematical problem solving, students at upper-elementary levels do not still have the required aptitudes to approach mathematical problems efficiently. They identify three main sources, or problem–solving components, that may cause these deficiencies: first, lack of domain-specific knowledge and skills; second, deficits in the heuristic, metacognitive, and affective aspects of mathematical competence; and last, inadequate domain-related beliefs about and attitudes towards mathematics and problem solving. Concerning student´s beliefs about mathematics, one of the most important components is the ability to assess and make judgments about one´s own performance and to notice discrepancies between one's own real –actual- performance and perceived performance. The correspondence between one´s 'perceived' performance and one´s 'actual' performance is referred to as *calibration*, and this measure has become an important topic of research in the last decades (Bouffard, Vezeau, Roy, & Lengelé, 2011; Dinsmore & Parkinson 2013; García, Rodríguez et al., 2016; Hadwin & Webster, 2013). Literature shows that students of different ages commonly make inaccurate evaluations or judgments of their performance, showing a tendency towards over-confidence. The relevance of calibration mechanisms in mathematics has been substantially demonstrated in mathematics problem-solving. Specifically, students with higher calibration skills tend to perform more successfully than students with lower calibration skills, and this has been explained by the greater degree of control over problem-solving processes in students with higher calibration abilities. However, inaccurate judgments of performance not only affect actual performance, but also motivation, persistence and interest in the task, as previous studies have suggested (Hadwin and Webster, 2013; Jacobse & Harskamp, 2012; Lipko et al., 2009; Rinne & Mazzocco, 2014). Those students who are less confident may also feel that they unable to tackle the given problem, which could thereby hinder them from utilising their full knowledge and previously learnt strategies; on the other hand, however, over-confidence may lead to excessive mistakes, frustration, and a lack of motivation in the face of failure. Accordingly, the relevance of taking these two variables (actual and perceived performance) into consideration should not be overlooked by researchers.

Nowadays, there is a good deal of evidence supporting the importance of problem-solving processes for learning, particularly regarding actual and self-perceived performance in mathematics problem-solving. However, there are two important questions that have been left unresolved to date: 1) Is mathematical problem-solving a linear or recursive process? 2) Does the organization of this process affect actual and/or perceived performance? A linear process would imply that there is a hierarchy, and that problem-solving is a sequential process (Krawec, 2012; Montague, Wager, & Morgan, 2000). So obstacles in the first phase would cause failure in the other phases. Alternatively, if the process is recursive, this would imply that planning, execution and evaluation may occur at any time during performance, which means recognizing the potentially iterative nature of the problem-solving process (Boonen, 2015). While the response to this question seems to be clear in the case of, for example, writing composition (i.e., the writing process is a recursive cyclical process; Lei, 2008; Smet, Brand-Gruwel, Leijten, & Kirschner, 2014), previous empirical evidence on mathematical problem-solving has not allowed any conclusion to be reached in this sense.

A brief description of the most representative models on mathematics problem-solving is provided below. It is important to note at this point that, while there are some models that can be easily defined as linear or hierarchical, the possible iterative and recursive nature of the problem-solving processes seems not to be clear in most models.

Mathematical problem-solving models

Table 1 summarizes some of the most representative models on problem-solving. All of these models are based on different phases or sub-processes, and (to a certain extent) involve planning, execution, and evaluation mechanisms.

- Please insert Table 1 here-

Within this context, some classical and widely-known models are presented, such as those proposed by Polya (1981), Mayer (2003), or Bransford and Stein (1993), or Montague (2000). The problem-solving activities that comprise these models can be typically summarized in two main phases: 1) problem understanding and representation, and 2) solution development (Babakhania, 2011; Kim, 2015; Krawek, 2012). For these authors, successful problem-solving is not possible without first interpreting and representing the problem adequately. A proper interpretation and representation indicates that the problem solver has understood the problem and serves as a powerful tool to guide them towards the solution plan (Babakhania, 2011). These authors also highlight the need for evaluation or revision at the end of the process. This is important, since revision strategies have been shown to be a determining factor for successful problem-solving, and one of the activities with which students struggle the most while performing these tasks (Cleary & Chen, 2009; García, Betts et al., 2016; García, Rodríguez et al., 2016; Montague, Enders, & Dietz, 2011). However, based on this type of characterization, it is still not clear whether evaluation and revision mechanisms are present during the whole process.

Although it is not properly a problem-solving model, Zimmerman´s Model (2000, 2008) is relevant in this sense, as metacognitive and self-regulatory mechanisms are linked to successful problem-solving (Babakhania, 2015; Cleary & Chen, 2009; Desoete & Roeyers, 2003; García, Rodríguez et al., 2016; Swanson, 1990). The three main phases of Zimmerman´s model (i.e., Forethoughts, Performance, and Self-reflection) correspond to the phases of planning, execution, and evaluation, referring to those activities that occur before, during and after performing a task or learning. This model defines the Self-

Regulated Learning (SRL) as a cyclical process in which planning, execution and revision can occur at any time during the task. Hence, it recognizes the potential recursive or iterative nature of the process. In this same line, the model of mathematical modeling and applied problem solving of Blum and Niss (1991) must be highlighted. These authors describe a process that implies the following stages: creation of a mathematical model of the situation described in the problem; working within mathematics (drawing conclusions, calculating and checking examples, applying known mathematical methods and results as well as developing new ones, etc.) which leads to obtain a mathematical result; and last, re-translating the result into the real world (validation of the model). The authors also recognize the presence of recursion within this process, as they state that various modifications of the previous model can occur as a result of the validation stage (p. 35). There are also some models that highlight, aside from recursion or iteration, the importance of monitoring one´s progress during the whole problem-solving process (e.g. Pretz, Naples, & Sternberg, 2003), which implies that control and evaluation mechanisms, must be present in every phase of the process. This is the case of the model proposed by Verschaffel et al. (1999). The authors conceptualize this model as a genuine strategy consisting in five stages eight heuristics that are especially valuable in the first two stages of the model (Table 1). Conceived as the goal of a learning environment, the model is aimed at facilitating students to become aware of the different phases involved in a competent problem-solving process (awareness training), developing the ability to monitor and evaluate one´s performance during the different phases of the problem-solving process (self-regulation training), and gaining mastery in the use of eight heuristic strategies that are especially useful during the first two stages of building a mental representation of the problem and deciding how to solve the problem (heuristic strategy training) (p. 201). More recently, Boonen (2015) empirically found some signs of recursion in mathematics problem-solving, on this occasion with a sample of teachers implementing an intervention program to support non-routine mathematics word problem-solving in upper-elementary students. The author analyzed the problem-solving process followed by teachers during the intervention sessions, taking the phases shown in Table 1 as a framework. These findings contrast with some models and research-based programs developed to support word problem-solving, which assume that it is a linear and hierarchical process; providing interesting evidence about the presence of recursion in the process instead.

The present study

A review of the relevant literature indicates that additional evidence should be gathered on the orchestration of the problem-solving process. As students´ difficulties in problem-solving can occur at any phase, a good understanding of this process and its characteristics may provide interesting insights into the mechanisms responsible for these difficulties. Certain useful guides for teachers to better approach mathematics problem-solving might be provided through a proper understanding of what happens during this complex activity, as theory and daily practice should be intimately linked (Oonk, Verloop, & Gravemeijer, 2015). In order to contribute to this goal, the present study aims to answer the following questions:

- Do students demonstrate a linear process during mathematics problem-solving, or there is some recursion present in this process?
- Is students' actual and self-perceived performance (i.e. success vs. failure) related to linearity or recursion in the process?

It is expected that:

- Students will be mainly linear in their process, although some signs of recursion will be present during performance.
- Statistically significant differences in the process will be found in students with differing actual and self-perceived performance in the problems (i.e., success-failure). As linear models seem to be more extended than those that recognize the possible presence of recursion today, it is likely that students who show a linear process will be more successful and make more positive judgments of performance than those showing a tendency to recursion.

Materials and Methods

Participants

A sample of 524 fifth and sixth grade students took part in this study. Ages ranged between 10 and 13 years old (M = 10.99, SD = 0.72). This sample comprised 220 students from the fifth grade (42%) and 304 students from the sixth grade of elementary school (58%). A total of 260 students (49.6%) were female, and 264 (50.4%) were male.

Students were recruited from 12 state and private schools in Northern Spain. Sample selection was made through accessibility procedures. Students volunteered for the study and presented informed consent from families. Children with a diagnosis of special educational needs or severe learning disabilities were excluded from the study.

Measures

As in other similar studies (García, Rodríguez et al., 2016; García, Cueli, Rodríguez, Krawec, & González-Castro, 2015), the Triple Task Procedure in Mathematics-TTPM- was used as a measure of the mathematics problem-solving process. This procedure is a modification of the traditional Triple Task technique, widely used in the study of the processes involved in composition writing (Olive & Piolat, 2002; Piolat, Olive, & Kellog, 2005). Actual performance was obtained through the students´ written responses to the mathematics problems, and self-perceived performance was obtained by asking students to judge whether they thought that they had given a correct response to each problem included in TTPM procedure.

Mathematics problem-solving process

The TTPM consists of the performance of three tasks: a primary task that elicits the cognitive process under investigation (i.e. solving a mathematics problem in this case); a probe task, in which response times [RTs] are measured; and a categorization task, in which students verbalize or label the actions or thoughts that are interrupted by the probe. The second task is used as a control task to ensure that students stay engaged in the evaluation process (i.e., extremely long RTs may be indicative of loss of attention or poor understanding of the instructions given), while the first and third tasks provide information about student performance (i.e. product or final result) and the process, respectively.

The procedure is as follows: for each mathematics problem, students are asked to select the category that best represents the activities they are engaged in at different times during performance (García, Rodríguez et al., 2016). During problem-solving, students hear a one second tone played at random intervals of between 40 and 45 seconds. This time interval allows evidence of the process to be gathered while trying to reduce possible interference in the process flow. On responding to the tone (i.e. probe task), students are presented with a category system that shows different activities involved in problem solving. They have to select a category identifying their current process from one of eight different activities: Reading the problem, drawing or summarizing, recalling similar problems, thinking about a solution, calculating, writing a response, reviewing, and correcting mistakes. RTs to the tones are registered. Students are initially trained, by means of examples, how to recognize and relate these categories to their own problem-solving process.

TTPM uses directed introspection in this categorization phase, as students are asked to categorize their actions or thoughts according to a given category system. This system is based on Zimmerman´s Self-Regulated Learning model (SRL; Zimmerman, 2000, 2008), in combination with Bransford and Stein´s (1993) IDEAL model of problem-solving. From the combination of both models, a system with eight categories or sub-processes emerged (García, Rodríguez et al., 2016). These sub-processes are organized into three higher level categories, corresponding to the main SRL phases of Planning, Execution, and Evaluation (Zimmerman, 2000, 2008). Table 2 shows the category system proposed by the authors. An additional category ("other") has been included to reflect thoughts or activities unrelated to problem-solving performance, such as day-dreaming (e.g., "I'm thinking about what I'm going to do this afternoon").

- Please insert Table 2 here-

Given its design, this technique is suitable for examination of the linearity and recursion of the process from the viewpoint of the main SRL phases of planning, execution and revision; understanding "linearity" as the tendency to progress forwards and hierarchically through the process, and "recursion" as the tendency to move backwards. Continuity, or the tendency to stay in the same phase for a period of time, is also evidenced by this technique, and was measured in the present study. Figure 1 exemplifies the process showed by two different students. The figure represents the category choice made by the students in each moment, providing evidence of the temporary organization of the process, including transitions between phases (either recursion or linearity) and continuity.

- Please insert Figure 1 here-

Mathematics problems

Students performed two word-based mathematical problems during the TTPM. The problems were based on everyday situations taken from the book "Problem-solving and comprehension" (Whimbey & Lochhead, 1999).

Problem 1: Beatriz lends €700 to Susana. But Susana borrows €1500 from Esther and €300 from Juana. In addition, Juana owes Esther €300 and Beatriz €700. One day they meet at Beatriz's home to settle their debts. Who went back home with €1800 more than she brought?

Problem 2: Paula, Mari, and Juana have a total of 16 dogs, 3 of which are poodles, 6 are greyhounds, and the rest of them are German shepherds and Pekinese dogs. Juana does not like poodles and Pekinese dogs, but she has 4 hounds and 2 German shepherds, leading to a total of 6 dogs. Paula has a poodle and 2 more dogs, which are German shepherds. Mari has 3 Pekinese dogs and several dogs of other breeds. Which breeds, and how many dogs of each breed, does Mari have?

Actual and self-perceived performance

A measure of actual performance (correct - incorrect) was obtained based on students´ written responses to the two mathematics problems.

A measure of self-perceived performance was obtained after task completion, by asking students to judge whether they thought that they had given a correct response to each problem. This variable was scored as "success" when students provided a positive judgment of performance, and as "failure" when they did not. This measurement system was used in previous studies with upper-elementary students (García, Kroesbergen, Rodríguez, González-Castro, & Gonzalez-Pienda, 2015; García, Rodríguez et al., 2016).

Procedure

The study was conducted in accordance with the Helsinki Declaration of the World Medical Association (Williams 2008), which reflects the ethical principles for research involving humans. The assessment procedure was collectively administered during a regular mathematics class. The TTPM was presented through a module enabled on Moodle. Students accessed the platform from their personal computers. They were given a username and password in order to guarantee anonymity. Two trained examiners led the evaluations. Only two mathematical problems were used in the present study according to the design of the TTPM, which requires that training and TTPM administration are conducted during the same session. This same procedure has been used in previous studies in mathematics and writing composition (García, Cueli et al., 2015; García, Betts et al., 2016; García, Rodríguez et al., 2016: Torrance, Fidalgo, & Robledo, 2015).

The first part of the session consisted of explaining the evaluation process to students. Then, they were trained in the recognition of the different process categories. In order to determine the students´ accuracy in categorizing the activities involved in mathematical problem solving, a pilot test was conducted, using the example of a student of their age (Alex) thinking aloud while solving a mathematical problem. Students had to recognize and categorize the sample student's activity at 18 different time points during the process (2 items per category, including the category called "other"). The students' categorizations were subsequently compared with those of one of our expert raters. Mean agreement between students´ codes and those of the expert was high (mean Cohen´s κ =.89). This result is consistent with those from recent studies on writing composition (Rodríguez, Grünke, González-Castro, García, & Álvarez-García, 2015; Torrance et al., 2015), and indicates a high reliability of the coding process in the current sample. This training phase lasted about 15 minutes.

Once the system of categories was understood, the TTPM was administered as previously described. The TTPM lasted until students indicated they had completed the problem, by clicking on a "finish" button set up for that purpose. Once students clicked on the "finish" button, a box appeared on the computer screen displaying the following question: "Do you think that you have given the correct response to the problem?" This allowed an indicator of students´ self-perceived performance to be gathered. The same procedure was repeated for the second problem.

The TTPM was designed so that, regardless of each student's response speed and the time they started or finished each problem, the time intervals between probes were the same for all participants. As evaluations were collectively administrated, students were provided with headphones. However, no more than 20 students were evaluated simultaneously.

Both the materials and the evaluation procedure were designed so that students found them appealing and motivating. The category system was represented through text and graphic symbols simultaneously. Mathematical problems were displayed on the computer screen and also presented on paper. Students were able to use the paper to write whatever they needed with the condition that they had to write their answer on the paper when they finished each problem. Once data from the process were retrieved, and the mathematics problems were marked, students´ actual and self-perceived performance in the problems was established in terms of success (1) or failure (0).

As extreme RTs during the TTPM were not identified, data from all the students were included in the analyses.

Students´ responses during the TTPM were coded for further analyses as follows: the reported activities in the planning phase (i.e., from reading to thinking about solutions) were assigned the number 1; those within the execution phase (i.e., calculating, and writing a response) were coded as the number 2; and those related to revision mechanisms (i.e., reviewing, and correcting mistakes) were assigned the number 3. Students´ responses during the TTPM were recorded for a maximum of 20 different times. They had to label their actions according to the category systems describe previously. Each time represented a 40-45 second interval, matching the different probes presented. Based on these time intervals, students had a maximum of 15 minutes to solve each problem. A significant number of students completed the task before reaching probe number 20. Once evidence from the process was obtained, transitions within the three different phases were established.

There were 19 possible transitions within phases, providing a sequence like this (e.g. 112223123). In this sequence, a student is performing the problem during 9 probe times. Thus, eight transitions (including continuity) are made. The student starts the task by planning and stays in this phase for two probes; then makes a transition to the execution phase, where they stay for a longer period of time; after that, the student moves to the revision phase and starts a new cycle of planning-execution-revision. In order to quantify these transitions, the phase reported by students at a certain probe time was compared to the phase they reported at the immediately preceding probe. Two different possibilities arose: 1) that students stayed performing an activity within the same phase (e.g. 11- continuity within the planning phase); or 2) that students made a transition from one phase to another, either forwards (e.g. 12- transition from planning to execution- linearity) or backwards (e.g. 21- transition from execution to planning-recursion). The absolute frequencies of these transitions, related to staying in each phase and transitions within phases, were calculated and used as dependent variables in the statistical analyses.

The Excel counting function was used to calculate absolute frequencies of transitions. Absolute frequencies for each transition were then averaged over the 524 participants in order to perform statistical analyses with means. In addition, and given that frequencies may provide a more exact estimation of process organization, they were also reported and analyzed in the present study.

Data analysis

As in previous studies, these analyses were initially based on mean frequency counts (García, Cueli et al., 2015; García, Betts et al., 2016; García, Rodríguez et al., 2016; Torrance et al., 2015). The dependent variables considered were the mean frequency with which students stayed focused on activities within the Planning, Execution and Revision phases while solving the problems, and the number of transitions they made from one phase to another (either forward or backward). In accordance with the objectives of the study, the data was analyzed in two steps.

Firstly, in order to examine the processes used by the students, and whether they were linear or recursive, means and standard deviations of the different dependent variables were reported. As dependent variables did not follow a normal distribution, non-parametric analyses were conducted. At a descriptive level, means and absolute frequencies both provided initial evidence about students´ tendency to continuity, recursion or linearity (Table 3).

Secondly, the Mann Whitney U-test was applied to analyze differences in the process between students with differing actual and self-perceived performance in the mathematics problems. Cliff's delta (δ) was used as a measure of effect size (Macbeth, Razumiejczyk, & Ledesma, 2011). This statistic provides a measure of dominance, or the degree of overlap between two distributions of scores. The value of this statistic ranges from –1 (if scores in Group 2 are larger than scores in Group 1) to +1 (if scores in Group 2 are smaller than scores in Group 1), and takes the value of zero if the two distributions are similar (i.e., absence of statistically significant group differences in the measured variables). Cohen (1988) established a bridge between Cohen's d and Cliff′s δ statistic: a δ value of .147 has an effect size of d = .20 (small effect); a δ value of .330 corresponds to an effect size of d = .50 (medium effect); and a δ of .474 has an effect size of d = .80 (large effect). Absolute frequencies of transitions were reported for each each group, as they provide a purer measure of the transition pattern than means. Based on these frequencies, odds ratios were calculated as an additional estimation of effect size (Tables 4 and 5). This statistic was used as an estimation of students´ likelihood to make a transition in comparison the total number of possible transitions. The higher the Odds ratio, the greater the probability of making this transition. The values of this statistic range from 0 to 1.

Separate analyses were conducted for each mathematics problem. SPSS v.23 (Arbuckle, 2010) was used to carry out the statistical analyses, with the exception of the non-parametric effect size analysis (i.e., Cliff´s delta statistic), for which Cliff's Delta Calculator (CDC: Macbeth et al., 2011) was used. A p-value ≤ .05 was established as the criterion of statistical significance.

Results

Mathematics Problem Solving Process: Linearity vs. Recursion

Table 3 shows descriptive statistics for each dependent variable. The first three transitions presented in the table are indicative of recursion, while the last three are related to linearity in the process. The second group of variables are indicative of continuity. As can be observed, the frequency of the different transitions is generally low, as students show a high tendency to stay in the same phase (high continuity), particularly in planning and execution phases. This pattern is even more visible when absolute frequencies are considered.

- Please insert Table 3 here-

In addition to this tendency to continuity, there are some transitions that are relatively frequent. The transition from planning to execution (related to the linearity of the process) is considerably more frequent than any other transition. However, the least frequently reported transitions occur from revision to planning, and from planning to revision in both mathematical problems. The first transition would be indicative of starting a new cycle in the process, while the second one would reflect the presence of monitoring activities from the initial phases of the process. Transitions from execution to revision (i.e., linearity), and from execution to planning (i.e., recursion) are also reported by students in some cases, with a similar frequency.

Relationship between Process and Actual Performance

Students´ actual and self-perceived performance in the two mathematical problems was measured in order to analyze the relationship between these two variables and the problem-solving process. For the first problem, 156 students (29.8%) were successful, while 368 (70.2%) answered incorrectly. A total of 449

(85.7%) gave a judgment of success about their performance in this problem, while 75 (14.3%) judged their performance as failure. For the second problem, 185 students (35.3%) solved it correctly, while 339 (64.7%) did not. A total of 403 (76.9%) judged their performance as successful in this problem, while 121 (23.1%) made a judgment of failure. This pattern indicates that there is a high mismatch between students´ perception of performance and actual performance.

Table 4 shows absolute frequencies for each dependent variable in the groups with differing actual performance in the mathematics problems. Figure 2, based on frequency means, shows the same tendency observed in the general group, with the activities within the phases of planning and execution being the most frequently reported by students. The transition from planning to execution (i.e. linearity) remains the most frequent observed category.

- Please insert Table 4 here-

Regarding inter-subject differences in the process as a function of students´ actual performance, statistical analyses showed that there is an absence of statistically significant differences between the students who solved the problem correctly and those who did not in Problem 1. There are however some statistically significant differences in Problem 2, specifically in the execution phase, $U = 25665.000$, $p < .001$, $\delta = .181$, as well as in transitions from planning to execution, $U = 28484.500$, $p = .033$, $\delta = .162$, and execution to revision, $U = 28347.000$, $p = .027$, $\delta = .112$; the latter two relate to the linearity of the process. The means indicate that students who successfully solved the second problem spent significantly more time executing. This group also reported planning before – and reviewing after- task execution (both related to the linearity of the process) to a greater extent than their peers with low performance in this problem (see Figure 2). As the δ statistic indicates, the effect size of these differences was low. The positive sign of the statistic confirms the direction of the differences, with higher means in the group who correctly judged their performance. Odds ratios in Table 4 also confirmed the tendency found regarding differences between groups, showing a higher likelihood of making the aforementioned transitions on this group. The Odds ratio values were very low in general, and only slightly higher in the case of the execution phase. This is explained by the generally low frequency of transitions that students made between the different phases, as they tended to continuity.

-Please include Figure 2 here-

Relation between Process and Self-Perceived Performance

Table 5 shows absolute frequencies for each transition between phases in the groups with differing self-perceptions of performance (success vs. failure) in the mathematics problems. Figure 3 shows frequency means of the transitions between the different phases in both groups. Again, a tendency towards continuity was found, and students spent most of their time in the planning and execution phases. In terms of linearity, the transition from planning to execution was the most frequent transition in both groups; while in the case of recursion, the transition from execution to planning was one of the most frequently reported by the students.

-Please insert Table 5 here-

Statistical analyses revealed that there were significant differences between the groups with differing judgments of performance in both problems. These differences were more notable in Problem 2. On the one hand, students who had a positive judgment about their performance in Problem 1 significantly differed from those that gave a judgment of failure in the following variables: transition from execution to planning, $U = 14217.500$, $p = .009$, $\delta = -.155$, and execution to revision, $U = 12501.500$, $p < .001$, $\delta = .257$, as well as in the amount of time they spent performing activities within the planning phase, $U = 11644.500$, $p < .001$, $\delta = -.308$. Students who believed they were correct in solving the first problem stayed in planning significantly less, and made the transition from execution to planning (i.e., recursion) to a lesser extent than the other group, while reviewing after execution (i.e., linearity) more frequently.

On the other hand, in Problem 2, group differences were found in a wider set of variables. Specifically, statistically significant differences were found in the amount of time that students stay in the phases of planning, $U = 18119.000$, $p < .001$, $\delta = -.256$, and execution, $U = 19425.000$, $p = .001$, $\delta = .203$, as well as in the frequency of transitions from revision to planning, $U = 23190.000$, $p = .005$, $\delta = -.109$, planning to execution, $U = 20511.500$, $p = .001$, $\delta = .158$, and execution to revision, $U = 20226.500$, $p = .001$, $\delta = .170$. Students who made a judgment of success in this second problem spent significantly less time in planning, while spending more time executing the task, in comparison to those who made a judgment of failure. The success group also went back from revision to planning (i.e., recursion) less frequently, but planned before- and reviewed after- execution (i.e., linearity) more frequently than their peers in the failure group. The effect size of these differences was generally low, although higher than in the actual performance analysis. The highest effect sizes were found in planning in both problems, and in the transition from execution to revision in Problem 1, with values close to a medium effect size. The sign of the δ statistic indicates the direction of the differences, the negative sign being indicative of higher means in the failure group. Odds ratios in Table 5 showed the same direction of means and effect sizes. The highest values of this statistic were found in the case of continuity in planning, which indicated a higher likelihood of the students making this transition as opposed to the others. Greater differences in Odds ratios were found in the transitions that turned out to be statistically significant in previous non-parametric analyses. The values of this statistic were very low, with the exception of continuity in the planning phase, as in previous analyses of actual performance.

-Please include Figure 3 here-

Discussion

This study aimed to analyze the mathematical problem-solving process in a sample of upper-elementary students in terms of linearity and recursion, and how this organization relates to students actual and self-perceived performance in these tasks. While the linear nature of the problem-solving process seems to be included in most of the previous models, some current research suggests that a form of recursion is also possible.

Firstly, although some examples of recursion were observed in this study (i.e. retrogressing from execution to planning), students used noticeably more linear than recursive processes. This is consistent with a characterization of the problem-solving process as a series of hierarchically organized activities, where students progress towards the solution in a linear and ordered function. This process is different from that found in the case of writing-composition activities, which have been demonstrated to show a cyclical nature (Lei, 2008; Smet, Brand-Gruwel, Leijten, & Kirschner, 2014). It is also important to note that the process demonstrated by this sample of students can be categorized as predominately continuity. Once students get involved in activities within a phase, they stay in that phase for a long period of time.

This pattern is especially evident in the planning and execution phases, but not in revision. Revision mechanisms are very scarce, and when present, they seem to be sporadic. This is coherent with previous research indicating that revision is an activity that many students overlook or even dislike, which commonly leads to poor performance in tasks (Cleary & Chen, 2009; García, Betts et al., 2016; Montague et al., 2011; Pennequin, Sorel, Nanty, & Fontaine, 2010). In fact, success rates in the present sample of students were very low for both problems.

Secondly, results showed a weak but statistically significant relationship between the organization of the process and students´ actual performance, but only in the second problem. It is worth noting in this case that these differences were found in the frequency of transitions related to linearity. Thus, no effect of recursion on actual performance was found. In this sense, success in the task was related to more time spent executing the task, the use of planning strategies before execution, and the presence of revision mechanisms once the task had been performed (i.e. a solution was given). This is important since, despite students not showing a great preference for revision in this study, the presence of these mechanisms has been shown to be significantly linked to good performance in the tasks. This finding is supported by previous literature (Cleary & Chen, 2009; García, Betts et al., 2016; Montague et al., 2011) recognizing the important role of planning and evaluation strategies for problem-solving in different educational stages. Thus, an important emphasis must be made in instructional programs to encourage students to put both mechanisms into practice.

Thirdly, regarding the relationship between process organization and students´ self-perceived performance, results indicated that students who judged their performance as correct demonstrated a more linear process (e.g. transition from execution to revision), while those who made a judgment of failure showed more signs of recursion (e.g. transition from execution to planning). Thus, linearity relates to good perceptions of one´s performance, while the presence of recursion is perceived by students as negative for performance. This finding is coherent with the current scientific scenario in which most intervention programs seem to be based on the linearity of the process (see Kim, 2015). Since most students are commonly encouraged to follow a structured, hierarchical, and linear path, it is unsurprising that they feel less confident when changes to this sequence are experienced. This might be perceived as negative by students, leading to a perception of failure. Another explanation may be that students may have found it to be a very difficult task, one for which they could not find an adequate strategy/solutions, which led to both recursive paths and to negative feelings about their performance in the problem. It is also worth noting that a large number of students in the current sample made a positive judgment of performance, which contrasts to the low levels of success rates in solving the tasks. Thus, there is a high mismatch between students´ actual and self-perceived performance. This is a tendency which has been widely reported in previous literature. Using the concept of calibration, or the degree of correspondence between one´s judgment of performance and actual performance (Hacker et al., 2008), there is a good deal of evidence that students in upper-elementary school commonly make inaccurate judgments, with a tendency to over-confidence in mathematical problem-solving and other areas (Bouffard et al. 2011; Dinsmore & Parkinson 2013; Hadwin & Webster 2013), which is indicative of poor metacognitive and self-regulatory skills. It is also important to consider the sort of problems used. Students commonly prefer problems that are directly solved through algorithmic methods, rather than those that imply heuristic methods, where effort has to be made “searching” for the solution (Căprioară, 2015). This second group of problems would correspond to the problems used in the present study, which imply understanding, planning and monitoring at a deep level in order to be successfully solved. Thus, it might be possible that the low rates of success could be due in part to the type of problems involved in the present study.

Conclusions

Solving mathematical problems is a difficult cognitive activity in many cases. This difficulty frequently comes from the complex nature and organization of the processes involved. Taking into account the cognitive and metacognitive nature of the process, the study of its organization can provide interesting insights into the strategies students use and whether they follow a structured, smooth path towards the solution or, on the contrary, it can provide information about the sub-process or phases they struggle with (Bonner, 2013).

Findings from this study demonstrated certain characteristics in the way students faced the task, such the use of as a linear -step by step- approach to the solution, poor revision strategies, or a perception of failure when they had to look back over the process and look at a new way to tackle the problem. This is important since this pattern is accompanied by low rates of success in problem-solving in the present sample, which suggests that some changes should be made in this context.

As mentioned before, obstacles in problem-solving may occur in any phase. Therefore, control and monitoring components should play a more important role throughout the entire process. In the present study, for instance, the students reported to using revision mechanisms after performance, but not after planning. Thus, it would seem that students do not tend to check their planning strategies. Understanding and representation of a problem usually occur during the planning phase and, in many cases, form the basis for subsequent development of a suitable method to solve the task (Babakhania, 2011; Kim, 2015; Krawek, 2012). Thus, correct monitoring during this phase may lead to essential steps in the progression towards a successful problem-solving strategy.

On the other hand, it is also worth noting that only a negligible proportion of students made a transition from revision to planning, which involves completing one cycle and starting another. This contrasts with previous conceptualizations of Self-Regulated Learning as being a cyclical and active process, as per those proposed by Zimmerman (2000; 2008).

Within this perspective, it is important for students to understand that solving a mathematical problem should be seen as a flexible and adaptive process, where they can move forwards and backwards in order to find the best solution path. In this sense, it is becoming more and more necessary to look at learning and teaching practices (Arslan & Yazgan, 2015; De Bock et al., 1998; Verschaffel et al., 1999). The inclusion of different dynamics, such us group work, modeling, or recursive prompts would greatly benefit students´ adoption of the correct process in each case. Thus, more effort should be made in the study of the nature and organization of mathematics problem-solving. Once good information about this is available to educators and the scientific community, more students will benefit from better adapted and maybe more efficient instructional approaches.

Limitations and future lines of research

Some limitations in the present study must be acknowledged: firstly, the use of only two mathematical problems. Including more tasks would permit gathering a clearer and more continuous measure of both actual and self-perceived performance (expressed in dichotomous terms in the present study).
Secondly, another possible limitation in the present study is that the TTPM may be intrusive for students, leading to a sort of "reactivity" (Bowles & Leow, 2005). This is a common factor in on-line measurements, and has been widely studied in the context of Thinking-aloud and traditional Triple Task

procedures, with the conclusion that, although the use of these measures may increase time taken to complete the task, simply instructing participants to verbalize their thoughts during a task does not alter the sequence of their cognitive processes or task performance (Bannert & Mengelkamp, 2008; Fox, Ericsson, & Best, 2011; Olive & Piolat, 2002). Nevertheless, the study of the process would greatly benefit from the adoption of new perspectives and methods, such as the use of eye-tracking systems (Van Viersen, Slot, Kroesbergen, Van't Noordende, & Leseman, 2013), which may provide interesting insights into student strategy use while minimizing reactance. The use of additional assessment techniques and instruments, such as cognitive interviews or observation, would also help improve the knowledge of such processes. Third, it is important to note that some additional variables, such as task complexity and familiarity, and the student´s general mathematics ability, must be considered as modulating variables for future studies. Both variables can influence the linear/recursive nature of the problem-solving process. Also, the type of instruction on mathematics problem solving received by the student must be considered (Pelaez, Cueli, Areces, García, & Rodríguez, 2017). Authors such as De Bock et al. (1998) suggest in this sense that, different aspects of the current culture and practice of school mathematics may develop in students a tendency to use linear models also in situations in which they are not applicable. Thus, the influence of these three components (students' familiarity with the task, students´ general mathematical ability, and school learning and teaching practices) and their possible interaction must be properly analyzed in further studies. Exploring these variables and their relation to the problem-solving process is also important taking into account the great number of students who did not solve the problems successfully. Lastly, the low to medium effect sizes found in this study indicate the need for caution about the scope of the findings. This could be partially explained by the high inter-subject variability observed in the problem-solving process (i.e. extremely high values of variance, skewness and kurtosis were found). It would be interesting to establish more homogeneous groups, maybe on the basis of additional components such as affective-motivational variables or academic achievement in the subject.

In summary, our contribution can be considered as a first attempt to answer the question about the linear or recursive nature of the mathematical problem-solving process. This is a question that has been raised for years, capturing the attention of theorist in Mathematics education. However, it seems to have no answer yet. Through the present study, the authors sought to integrate a great part of the information gathered for decades at a theoretical level, while pointing out the need to propose and test empirical data-analysis methods that allow us to validate some of the theoretical models proposed up to the date, and therefore, be closer to answering the question about whether the order of the phases in solving mathematical problems matters.

References

Arbuckle, J. L. (2013). *SPSS (version 22.0) [Computer program]*. Chicago: SPSS.

Arslan, C., & Yazgan, Y. (2015). Common and Flexible Use of Mathematical Non Routine Problem Solving Strategies. *American Journal of Educational Research, 3(12),* 1519-1523. doi:10.12691/education-3-12-6.

Babakhania, N. (2011). The effect of teaching the cognitive and meta-cognitive strategies (self-instruction procedure) on verbal math problem-solving performance of primary school students with verbal problem- solving difficulties. *Procedia - Social and Behavioral Sciences, 15,* 563–570.

Bannert, M., & Mengelkamp, C. (2008). Assessment of metacognitive skills by means of instruction to think aloud and reflect when prompted. Does the verbalization method affect learning? *Metacognition and Learning, 3*, 39–58.

Blum, W., & Niss, M. (1991). Applied mathematical problem solving, modelling, applications, and links to other subjects - state, trends and issues in mathematics instruction. *Educational Studies in Mathematics, 22*, 37-68.

Bonner, S.M. (2013). Mathematics Strategy Use in Solving Test Items in Varied Formats. *The Journal of Experimental Education, 81*(3), 409-428

Boonen, A. J. H. (2015). *Comprehend, visualize & calculate: Solving mathematical word problems in contemporary math education. Dissertation Thesis.*

Bouffard, T., Vezeau, C., Roy, M., & Lengelé, A. (2011). Stability of biases in self-evaluation and relations to well-being among elementary school children. *International Journal of Educational Research, 50,* 221–229. doi:10.1016/j.ijer.2011.08.003.

Bowles, M. A., & Leow, R. P. (2005). Reactivity and type of verbal report in SLA research methodology. *Studies in Second Language Acquisition, 27*, 415–440.

Bransford, J. D., & Stein, B. S. (1993). *The ideal problem solver: A guide for improving thinking, learning and creativity (2nd ed.).* New York: W.H. Freeman.

Căprioarăa, D. (2015). Problem solving - purpose and means of learning mathematics in school. *Procedia - Social and Behavioral Sciences, 191,* 1859 – 1864. doi:10.1016/j.sbspro.2015.04.332.

Carlson, M. P., & Bloom, I. (2005). The cyclic nature of problem solving: an emergent multidimensional problem solving framework. *Educational Studies in Mathematics, 58*(1), 45-75. doi:10.1007/s10649-005-0808-x

Cleary, T. J., & Chen, P. (2009). Self-regulation, motivation, and math achievement in middle school: variations across grade level and math context. *Journal of School Psychology, 47,* 291–314. doi:10.1016/j.jsp.2009.04.002.

De Bock, D., Verschaffel, L., & Janssens, D. (1998). The predominance of the linear model in secondary school students´ solutions of word problems involving length and area of similar plane figures. *Education Studies in Mathematics, 35*, 65-83.

De Corte, E., & Somers, R. (1982). Estimating the outcome of a task as a heuristic stratege in arithmetic problem solving: A teacher experiment with sixth-graders. *Human Learning, 1,* 105-121.

Desoete, A., & Roeyers, H. (2003). Can off-line metacognition enhance mathematical problem solving? *Journal of Educational Psychology, 9*5, 188–200.

Dinsmore, D. L., & Parkinson, M. M. (2013). What are confidence judgments made of? Students' explanations for their confidence ratings and what that means for calibration. *Learning and Instruction, 24*, 4–14. doi:10.1016/j.learninstruc.2012.06.001.

Dunlosky, J., & Rawson, K. A. (2012). Over-confidence produces underachievement: inaccurate self-evaluations undermine students' learning and retention. *Learning and Instruction, 22,* 271–280. doi:10.1016/j.learninstruc.2011.08.003.

Dunlosky, J., & Thiede, K. W. (2013). Four cornerstones of calibration research: why understanding students' judgments can improve their achievement. *Learning and Instruction*, *24,* 58–61. doi:10.1016/j.learninstruc.2012.05.002.

Finn, B., & Metcalfe, J. (2014). Over-confidence in children's multi-trial judgments of learning. *Learning and Instruction, 32*, 1–9. doi:10.1016/j.learninstruc.2014.01.001.

Fox, M. C., Ericsson, K. A., & Best, R. (2011). Do procedures for verbal reporting of thinking have to be reactive? A meta-analysis and recommendations for best reporting methods. *Psychological Bulletin, 137,* 316–344. doi:10.1037/a0021663.

García, T., Betts, L., González-Castro, P., González-Pienda, J. A., & Rodríguez, C. (2016). On-line assessment of the process involved in Maths problem-solving in fifth and sixth grade students: Self-regulation and achievement. *Revista Latinoamericana de Investigación en Matemática Educativa, 19*(2), 165-186. doi: 10.12802/relime.13.1922

García, T., Cueli, M., Rodríguez, C., Krawec, J., & González-Castro, P. (2015). Metacognitive knowledge and skills in students with deep approach to learning. Evidence from mathematical problem solving. *Journal of Psychodidactics, 20*(2), 209-226. doi: 10.1387/RevPsicodidact.13060

García, T., Kroesbergen, E.H., Rodríguez, C., González-Castro, P., & Gonzalez-Pienda, J.A. (2015). Factors involved in making post-performance judgments in mathematics problem solving. *Psicothema, 27*(4), 374-380. doi: 10.7334/psicothema2015.25

García, T., Rodríguez, C., González-Castro, P., González-Pienda, J. A., & Torrance, M. (2016). Elementary students’ metacognitive processes and post-performance calibration on mathematical problem-solving tasks. *Metacognition and Learning, 11,* 139–170. doi: 10.1007/s11409-015-9139-1

Hacker, D. J., Bol, L., & Bahbahani, K. (2008). Explaining calibration accuracy in classroom contexts: the effects of incentives, reflection, and explanatory style. *Metacognition and Learning, 3,* 101–121. doi:10.1007/s11409-008-9021-5.

Hadwin, A. F., & Webster, E. A. (2013). Calibration in goal setting: examining the nature of judgments of confidence. *Learning and Instruction, 24*(12), 37–47. doi:10.1016/j.learninstruc.2012.10.001.

Jacobse, A. E., & Harskamp, E. G. (2012). Towards efficient measurement of metacognition in mathematical problem-solving. *Metacognition and Learning, 7*(2), 133–149

Jitendra, A. K., Dupuis, D. N., & Zaslofsky, A. F. (2014). Curriculum-Based Measurement and Standards-Based Mathematics: Monitoring the Arithmetic Word Problem-Solving Performance of Third-Grade Students at Risk for Mathematics Difficulties. *Learning Disability Quarterly, 2,* 1-11. doi: 10.1177/0731948713516766

Kim, M. K (2015). Models of learning progress in solving complex problems: Expertise development in teaching and learning. *Contemporary Educational Psychology, 42*, 1-16.

Krawec, J. L. (2012). Problem representation and mathematical problem solving of students of varying math ability. *Journal of Learning Disabilities, X*, 1-13. doi:10.1177/0022219412436976

Laua, C., Kitsantasb, A., & Millerc, A. (2015). Using microanalysis to examine how elementary students selfregulate in math: A case study. *Procedia - Social and Behavioral Sciences*, *174*, 2226 – 2233. doi: 10.1016/j.sbspro.2015.01.879

Lei, X. (2008). Exploring a sociocultural approach to writing strategy research: Mediated actions in writing activities. *Journal of Second Language Writing, 17*, 217–236.

Lipko, A. R., Dunlosky, J., Hartwig, M. K., Rawson, K. A., Swan, K., & Cook, D. (2009). Using standards to improve middle school students' accuracy at evaluating the quality of their recall. *Journal of Experimental Psychology Applied, 15*(4), 307–318. doi:10.1037/a0017599

Macbeth, G., Razumiejczyk, E., & Ledesma, R. D. (2011). Cliff's delta calculator: a non-parametric effect size program for two groups of observations. *Universitas Psychologica*, *10*(2), 545–555.

Mayer, R. E. (2003). *Mathematical problem solving. In J. M. Royer (ed.), Mathematical Cognition.* Greenwich, CT: Info age Publishing

Montague, M., Warger, C., & Morgan, T. H. (2000). Solve It! Strategy instruction to improve mathematical problem solving. *Learning Disabilities Research & Practice, 15*, 110-116. doi:10.1207/SLDRP1502_7

Montague, M., Enders, G., & Dietz, S. (2011). Effects of cognitive strategy instruction on math problem-solving of middle school students with learning disabilities. *Learning Disability Quarterly, 34*(4), 262–272. doi:10.1177/073i9487M421762.

Olive, T., & Piolat, A. (2002). Suppressing visual feedback in written composition: effects on processing demands and coordination of the writing processes. *International Journal of Psychology, 37*(4), 209–218. doi:10.1080/00207590244000089

Oonk, W., Verloop, N., & Gravemeijer, K. P. E. (2015). Enriching Practical Knowledge: Exploring Student Teachers' Competence in Integrating Theory and Practice of Mathematics Teaching. *Journal for Research in Mathematics Education, 46*(5), 559-598.

Pelaez, N., Cueli, M., Areces, D., García, T., & Rodríguez, C. (2017). Effect of Integrated Dynamic Representation on Mathematical Competence and Care in Children with Attention Deficit with Hyperactivity Disorder. *Journal of Psychology and Education, 12*(2), 105-115. Doi: https://doi.org/10.23923/rpye2017.12.149

Pennequin, V., Sorel, O., Nanty, I., & Fontaine, R. (2010). Metacognition and low achievement in mathematics: the effect of training in the use of metacognitive skills to solve mathematical word problems. *Thinking and Reasoning, 16*(3), 198–220. doi:10.1080/13546783.2010.509052.

Piolat, A., Olive, T., & Kellogg, R. T. (2005). Cognitive effort during note taking. *Applied Cognitive Psychology, 19*, 291–312. doi:10.1002/acp.1086.

Polya, G. (1981) *Mathematical Discovery: On Understanding, Learning and Teaching Problem Solving.* New York: Wile

Pretz, J. E., Naples, A. J., & Sternberg, R. J. (2003). *Recognizing defining, and representing problems.* In J. E. Davidson & R. J. Sternberg (Eds.), The psychology of problem solving (pp. 3–30). New York: Cambridge University Press.

Raynal, F., & Rieunier, A. (1997). *Pédagogie: Dictionnaire des concepts clés.* Paris: ESF, éditeur.

Rinne, L. F., & Mazzocco, M. M. M. (2014). Knowing right from wrong in mental arithmetic judgments: calibration of confidence predicts the development of accuracy. *PLoS ONE, 9*(7), e98663. doi:10.1371/journal.pone.0098663

Rodríguez, C., Grünke, M., González-Castro, P., García, T., & Álvarez-García, D. (2015). How Do Students With Attention-Deficit/Hyperactivity Disorders and Writing Learning Disabilities Differ From Their Nonlabeled Peers in the Ability to Compose Texts? *Learning Disabilities: A Contemporary Journal 13*(2), 157-175.

Schoenfeld, A. H. (1992). Learning to think mathematically: Problem solving, metacognition, and sense-making in mathematics. In D. Grouws (Ed.), *Handbook for Research on Mathematics Teaching and Learning* (pp. 334-370). New York: MacMillan.

Silver, E. A., Ghousseini, H., Gosen, D., Charalambous, C., & Font Strawhun, B. T. (2005). Moving from rhetoric to praxis: Issues faced by teachers in having students consider multiple solutions for problems in the mathematics classroom. *Journal of Mathematical Behavior, 24*, 287–301.

Smet, M. J. R., Brand-Gruwel, S., Leijten, M., & Kirschner, K. (2014). Electronic outlining as a writing strategy: Effects on students´writing products, mental effort and writing process. *Computers and Education, 78*, 352-366.

Swanson, H. L. (1990). Influence of metacognitive knowledge on problem solving. *Journal of Educational Psychology, 82,* 306–314.

Tambychika, T., & Mohd Meerahb, T. S. (2010). Students' Difficulties in Mathematics Problem-Solving: What do they Say? *Procedia - Social and Behavioral Sciences, 8,* 142–151. doi:10.1016/j.sbspro.2010.12.020

Torrance, M., Fidalgo, R., García, J. N. (2007). The teachability and effectiveness of cognitive self-regularion in sixth grade writers. *Learning and Instruction, 17*(3), 265-285.

Torrance, M., Fidalgo, R., & Robledo, P. (2015). Do sixth-grade writers need process strategies? *British Journal of Educational Psychology, 85*(1), 91–112. doi: 10.1111/bjep.12065

Van der Schoot, M., Bakker Arkema, A. H., Horsley, T. M., & Van Lieshout, E. D. C. M. (2009). The consistency effect depends on markedness in less successful but not successful problem solvers: An eye movement study in primary school children. *Contemporary Educational Psychology, 34*, 58–66. http://dx.doi.org/10.1016/ j.cedpsych.2008.07.002

Van Viersen, S., Slot, E. S., Kroesbergen, E. H., Van't Noordende, J. E., & Leseman. P. (2013). The added value of eye-tracking in diagnosing dyscalculia: a case study. *Frontiers in Psychology, 4*: 617. doi: 10.3389/fpsyg.2013.00679

Verschaffel, L., De Corte, E., Lasure, S., Van Vaerenbergh, G., Bogaerts, H., & Ratinckx, E. (1999). Learning to solve mathematical application problems: a design experiment with fifth graders. *Mathematical Thinking and Learning, 1,* 195-230.

Whimbey, A., & Lochhead, J. (1999). *Problem-solving and comprehension*. Hillsdale: Erlbaum.

Zimmerman, B. (2000). *Attaining self-regulation. A social cognitive perspective.* In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), Handbook of self-regulation (pp. 13–39). San Diego, CA: Academic.

Zimmerman, B. (2008). Investigating self-regulation and motivation: historical background, methodological developments, and future prospects. *American Educational Research Journal*, *45*(1), 166–183. doi:10.3102/0002831207312909.

Table 1. Problem-solving Models

Model	Phases
Linear (hierarchical) models	
Polya (1981)	- Understanding the problem - Planning - Performing the plan - Confirmation of the answer
IDEAL model (Bransford & Stein, 1993)	- Identification of the problem - Definition of the problem - Exploration of possible solutions - Acting according to the solution plan - Review of the last stages (Looking back)
Montague (2000)	- Read - Paraphrase - Visualize - Hypothesize - Estimate (predict the answer) - Compute - Check
Mayer (2003)	- Translation - Interpretation - Planning - Execution
Pretz et al. (2003)	- Recognizing the problem - Defining and interpreting the problem - Developing a solution strategy - Organizing one´s knowledge about the problem - Allocating mental resources to solve the problem - Monitoring one´s progress towards the goal - Evaluating the solution
Recursive models	
Verschaffel et al. (1999)	-Build a mental representation of the problem Heuristics: - Draw a picture - Make a list, scheme or a table - Distinguish relevant information from irrelevant data - Use your own real-world knowledge -Decide how to solve the problem Heuristics: - Make a flowchart - Guess and check - Look for a pattern - Simplify the numbers -Execute the necessary calculations -Interpret the outcome and formulate an answer -Evaluate the solution
*Zimmerman´s SRL Model (2000, 2008)	- Forethoughts - Performance - Self-reflection
Boonen (2015)	- Read the problem - Understand the text - Visualize the problem structure - Hypothesize a plan to solve the problem - Compute the required operation - Check your answer

Note. *Zimmerman´s Model is included as it is an example of the process involved in general learning and supposed the basis for different models based on problem-solving.

Table 2. Category System of problem-solving process (as a combination of Zimmerman and Bransford and Stein’s models)

SRL Model	IDEAL Model	Process categories (I am ...)
Planning	Identification of the problem	Reading
	Definition and representation	Drawing or summarizing
		Recalling similar problems
	Exploration of possible strategies	Thinking about a solution
Execution	Action based on the strategy	Calculating
		Writing a response
Revision	Look at effects of solutions	Reviewing
		Correcting mistakes
“Other”		Doing something unrelated

Note. Retrieved from García, Rodríguez et al. (2016).

Table 3. Descriptive statistics for each dependent variable (transitions within phases). Problems 1 and 2

Transitions	Problem 1			Problem 2		
	M	*SD*	*AF*	*M*	*SD*	*AF*
	Recursion					
rev-plan	.053	.241	28	.034	.211	18
rev-exec	.116	.385	61	.068	.295	36
exec-plan	.431	.716	226	.316	.604	166
TOTAL	.595	.898	314	.421	.762	219
	Continuity					
plan-plan	3.162	2.602	1657	2.774	2.864	1454
exec-exec	2.524	2.785	1323	2.509	2.470	1315
rev-rev	.319	.782	167	.200	.677	105
TOTAL	6.042	3.369	3147	5.501	2.951	2873
	Linearity					
plan-exec	1.162	.693	609	1.068	.638	560
exec-rev	.456	.576	239	.347	.503	182
plan-rev	.099	.311	52	.064	.275	34
TOTAL	1.711	.947	900	1.484	.862	775

Note. M = Mean frequency of transitions within phases; AF = Absolute frequencies; Rev = revision phase; Plan = planning phase; Exec = execution phase. Total sample = 524.

Table 4. Absolute Frequencies (AF) of the groups in the dependent variables (transitions within phases). Groups divided according to actual performance (correct vs incorrect). Problems 1 and 2

	Problem 1		Problem 2	
	Correct	Incorrect	Correct	Incorrect
Transitions	*AF(OR)*	*AF(OR)*	*AF(OR)*	*AF(OR)*
		Recursion		
rev-plan	8 (.002)	20 (.002)	4 (.001)	14 (.002)
rev-exec	16 (.005)	45 (.006)	12 (.003)	24 (.003)
exec-plan	56 (.018)	170 (.024)	57 (.016)	109 (.017)
		Continuity		
plan-plan	461 (.155)	1196 (.171)	466 (.132)	988 (.153)
exec-exec	374 (.126)	949 (.135)	532 (.152)	783 (.121)
rev-rev	45 (.015)	122 (.017)	46 (.013)	59 (.009)
		Linearity		
plan-exec	179 (.060)	430 (.061)	212 (.060)	348 (.054)
exec-rev	73 (.024)	166 (.023)	75 (.021)	107 (.016)
plan-rev	16 (.005)	36 (.005)	10 (.002)	24 (.003)

Note. OR = Odds Ratio; Plan = planning phase; Exec = execution phase; Rev = revision phase; Problem 1 Correct (n = 156); Problem 1 Incorrect (n = 368), Problem 2 Correct (n = 185), Problem 2 Incorrect (n = 339).

Table 5. Absolute Frequencies (AF) of the groups in the dependent variables (transitions within phases). Groups divided according to self-perceived performance (correct vs. incorrect). Problems 1 and 2

	Problem 1		Problem 2	
	Correct	Incorrect	Correct	Incorrect
Transitions	*AF(OR)*	*AF(OR)*	*AF(OR)*	*AF(OR)*
	Recursion			
rev-plan	23 (.002)	5 (.003)	2 (< .001)	10 (.004)
rev-exec	54 (.006)	7 (.005)	27 (.003)	9 (.004)
exec-plan	185 (.021)	41 (.029)	129 (.016)	37 (.016)
	Continuity			
plan-plan	1.317 (.154)	340 (.238)	988 (.129)	466 (.202)
exec-exec	1.078 (.126)	245 (.171)	1075 (.140)	240 (.104)
rev-rev	149 (.017)	18 (.012)	86 (.011)	19 (.008)
	Linearity			
plan-exec	521 (.061)	88 (.061)	450 (.058)	110 (.047)
exec-rev	221 (.026)	18 (.012)	154 (.020)	28 (.012)
plan-rev	47 (.005)	5 (.003)	25 (.003)	9 (.004)

Note. OR = Odds Ratio; Plan = planning phase; Exec = execution phase; Rev = revision phase; Problem 1 Correct (n = 449); Problem 1 Incorrect (n = 75), Problem 2 Correct (n = 403), Problem 2 Incorrect (n = 121).

Figure 1. Example of TTPM design. The figure shows the category choice made by two different students.

Action	Moment 1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
Reading		X														
Drawing or summarizing						X										
Recalling similar problems																
Thinking about a solution				●	X					●						
Calculating							X									
Writing a response	●	●	●	X	●	●	●	●	●		●	●				
Reviewing																
Correcting mistakes													●	●		
Doing something unrelated	X		X													

X Student 1

● Student 2

Figure 2. Mean frequency of transitions between phases (groups based on Actual Performance)

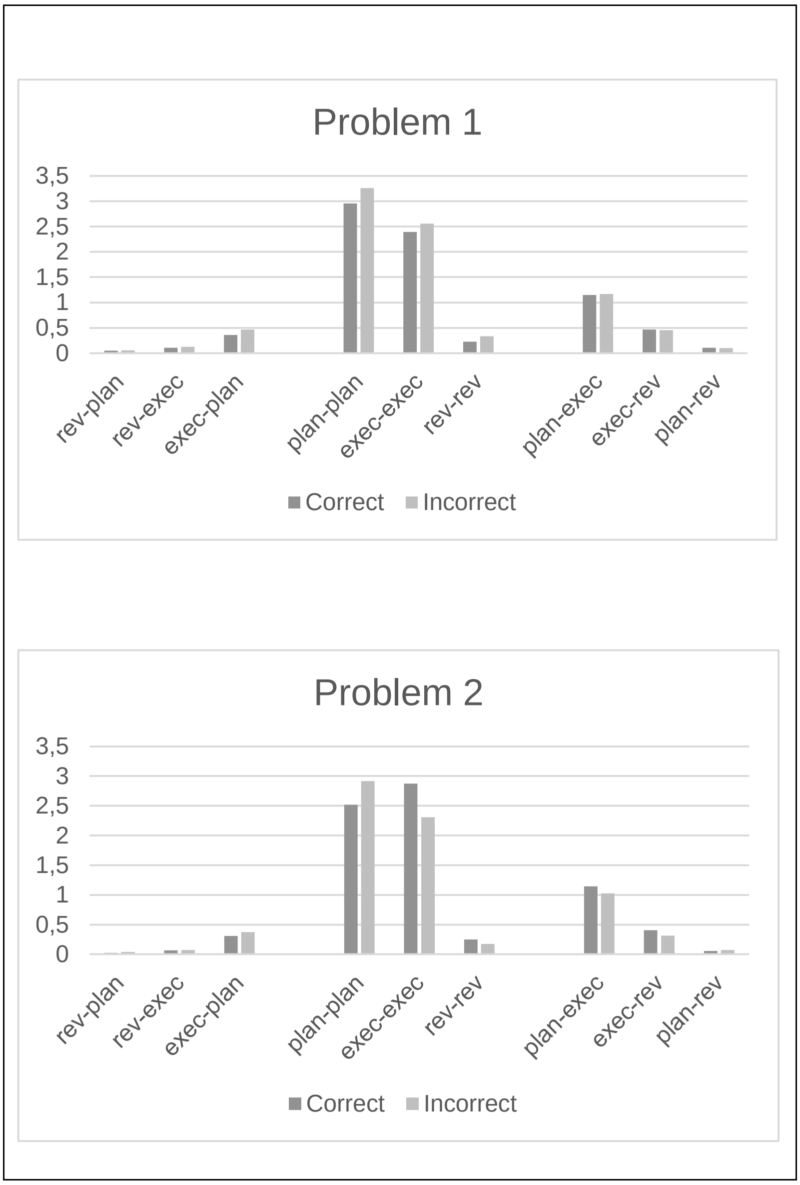

Figure 3. Mean frequency of transitions between phases (groups based on Self-Perceived Performance)