



Thematic Review

Long-term effects of metacognitive strategy instruction on student academic performance: A meta-analysis



Hester de Boer*, Anouk S. Donker¹, Danny D.N.M. Kostons, Greetje P.C. van der Werf

GION Education/Research, University of Groningen, Groningen, The Netherlands

ARTICLE INFO

Keywords:

Academic performance
Metacognitive strategy instruction interventions
Long-term effects
Meta-Analysis
Moderator analysis

ABSTRACT

Meta-analyses have shown the positive effects of strategy instruction on student performance; however, little meta-analytical research has been conducted on its long-term effects. We examined the long-term effects of 48 metacognitive strategy instruction interventions on student academic performance. The results show a very small increase of the effect at long-term compared with the posttest effects. The instruction effect at posttest increased from Hedges' $g = 0.50$ to 0.63 at follow-up test. Moderator analyses showed that low SES students benefited the most at long-term. Furthermore, instructions including the cognitive strategy 'rehearsal' had lower long-term effects compared to interventions without this component. Other specific strategies (within categories metacognitive, cognitive, management, or motivational) did not moderate the overall positive long-term effect of metacognitive strategy instructions. Particular attributes of the intervention – subject domain, measurement instrument, duration, time between posttest and follow-up test, and cooperation – neither had an impact on the follow-up effect.

1. Introduction

In response to the challenges of preparing for careers in the 21st century, students are required to become more and more in control of their own learning (OECD, 2014). Self-regulated learning is a way for students to effectively deal with such control. Pintrich (2000) described self-regulated learning as “an active, constructive process whereby learners set goals for their learning and then attempt to monitor, regulate, and control their cognition, motivation, and behavior, guided and constrained by their goals and the contextual features in the environment” (p. 453). A typical characteristic of self-regulated learners is that they use strategies which facilitate and enhance their learning process and consequently their academic performance (Zimmerman, 1986; 2002). This is why in education many interventions have been conducted to teach students learning strategies, which appeared to be effective from the results of multiple meta-analyses (Chiu, 1998; Dignath & Büttner, 2008; Donker, de Boer, Kostons, Dignath-van Ewijk, & van der Werf, 2014; Haller, Child, & Walberg, 1988; Hattie, Biggs, & Purdie, 1996). In the daily school context, students do not have to, and are not able to, fully regulate their own learning, as the teacher prescribes what to learn when, provides instruction, and assesses to what extent students master the learning goals. However, to prepare students for learning at a later stage, they should be instructed how to regulate and control their own learning by teaching students how to plan, monitor and evaluate their learning and provide them with a range of strategies that facilitate learning. These learning skills and the knowledge about when, how and which strategy to use form part of students' metacognition. Although the aim of teaching learning skills lies in transfer towards future learning tasks,

* Corresponding author. GION, University of Groningen, Grote Rozenstraat 3, 9712 TG Groningen, The Netherlands.

E-mail addresses: hester.de.boer@rug.nl (H. de Boer), anouk.donker@nhl.nl (A.S. Donker), d.d.n.m.kostons@rug.nl (D.D.N.M. Kostons), m.p.c.van.der.werf@rug.nl (G.P.C. van der Werf).

¹ Anouk Donker is currently working for the NHL Hogeschool Leeuwarden.

<https://doi.org/10.1016/j.edurev.2018.03.002>

Received 10 November 2017; Received in revised form 9 February 2018; Accepted 2 March 2018
Available online 21 March 2018

1747-938X/ © 2018 Elsevier Ltd. All rights reserved.

only a few researchers have looked at the long-term effectiveness of interventions that provided training in such strategies. The main purpose of this research paper is to provide a meta-analysis of these long-term effects. The prior meta-studies have already identified characteristics of effective instructions, which is a prerequisite for developing evidence-based effective instruction. But this knowledge is not sufficient. Knowledge about the long term effects of these strategy instructions is highly relevant and necessary too. We want to prepare students for a society which rapidly develops, and in which continuously new knowledge and skills are required. In order to properly prepare young people for such a society, education has to equip them with the necessary learning skills to keep them learning throughout their lives. Therefore, it is important to know whether instruction in learning skills has sustainable effects after the instruction has ended. It tells us to what extent we are actually able to prepare students for learning after the official school period. This information also affects educational practice: if students' academic performance drops after the strategy instruction has ended, this is an indication that strategy instruction should be addressed repeatedly. However, if the positive effects of strategy instruction sustain, it tells us that we do not have to implement strategy instruction in each lesson. Furthermore, knowledge about which factors affect the sustainability of the strategy instruction is necessary. Are short instructions just as effective as longer ones, does sustainability depend on the content of the instruction, or on the characteristics of the students who receive the instruction? All these aspects are important to know for the development of effective strategy instructions.

1.1. Learning strategy types

According to Alexander, Graham, and Harris (1998), the ability to apply a learning strategy is a form of procedural knowledge: the 'how to' knowledge. This 'how to' can be very diverse, as shown by the large number of strategies presented in the literature, ranging from simple re-reading strategies to more complex methods for synthesizing knowledge. A common characteristic, however, is that the strategy is applied intentionally and purposefully, with in mind the learning goal. The strategy is used as a means to reach the learning goal. Due to the plethora of possible strategies, various strategy-classification systems have been used in an attempt to organize these strategies. In the various classification systems, three main types of strategies recur: cognitive, metacognitive, and management strategies (e.g., Boekaerts, 1997; Mayer, 2008; Pressley, 2002; Weinstein & Mayer, 1986).

Cognitive strategies are strategies aimed at increasing the understanding and remembering of particular information; they make the material more meaningful for the learners. Such strategies should help in the development of coherent mental representations by imposing structure on the information gathered, which in turn enables students to integrate new information with existing knowledge more easily (Mayer, 2008). The cognitive strategies can be divided into several groups: elaboration strategies through which learners build connections between new material and what is already known (for example, through summaries, Brown & Palincsar, 1989; Cromley, Snyder-Hogan, & Luciw-Dubas, 2010); rehearsal strategies, which help learners store information in the memory by repeating and remembering its content; and organization strategies, which help learners to categorize and structure the information (Mayer, 2008; Pintrich, Smith, Garcia, & McKeachie, 1991).

Metacognitive strategies are methods which facilitate and regulate cognition. Because the use of these strategies involves monitoring and controlling one's own learning, including the application of cognitive strategies, metacognitive strategies are considered higher-order skills, and are more difficult to teach than cognitive strategies (Veenman, Van Hout-Wolters, & Afflerbach, 2006). Zimmerman (2002) distinguished three phases of metacognitive activities. During the first, the forethought phase, planning strategies are addressed: learners analyze the task and plan how to tackle it. Next, in the performance phase, the actual learning or task execution takes place. In this phase the monitoring strategies come into play; they are used to check the understanding of the material (Azevedo & Cromley, 2004; Kostons, Van Gog, & Paas, 2010). The last phase in the learning process is self-reflection, during which a learner evaluates the learning process or product, which in turn provides information for the planning of the next learning task.

Finally, management strategies are applied to deal with the context of the learning environment (Pintrich, 2000). If effectively used, they create the optimal surroundings for learning. These strategies can be aimed at the learner him/herself (Pintrich et al., 1991), such as persisting on a task in spite of facing difficulties (effort management) (Pintrich, 2004; Zimmerman, 1990), or at others such as peer students with whom students can cooperate (e.g., Palincsar & Brown, 1984), or at the physical environment, for example by finding a quiet place to work.

1.2. Metacognitive knowledge and motivation

Alexander et al. (1998) described the application of learning strategies as procedural knowledge, but in order to apply strategies effectively, students also need conditional knowledge on when to use these strategies, which is metacognitive knowledge. Flavell (1979) has defined metacognitive knowledge as knowledge about particular variables and how they act and interact to affect the course and outcome of cognitive enterprises. Metacognitive knowledge includes general knowledge about when to use which strategy and the extent to which a strategy is usually effective, and knowledge of oneself as a strategic learner (Pintrich, 2002). Students who lack metacognitive knowledge do not understand when or why to use these strategies (Pintrich, 2002). It is, therefore, of paramount importance that learning strategies are not simply taught as procedures to be followed, but that students are also taught when to use specific strategies. Students who have more metacognitive knowledge are more likely to successfully solve problems (Bransford, Brown, Cocking, & National Research Council, 1999; Schneider & Pressley, 1997; Weinstein & Mayer, 1986).

Furthermore, as learning is an active process, a learner also has to be motivated to use learning strategies. There are several relevant motivation aspects that affect how students engage in a learning task. We distinguished three aspects. The first is self-efficacy beliefs (Pintrich, 2003), which is the perception of one's ability to perform a task. The second aspect is task value (Wigfield & Eccles,

2002). This is the degree to which a task is valued as important or interesting. The third aspect we distinguished is goal orientation (Harackiewicz, Barron, Pintrich, Elliot, & Thrash, 2002); this refers to whether a student is intrinsically or extrinsically motivated to perform a learning task. These motivation aspects can be addressed during strategy instruction to facilitate the learning process.

1.3. Characteristics of effective strategy instruction

Strategy instruction helps students to apply learning strategies more effectively, and to become more successful learners. Several meta-analyses have been performed to investigate the effects of strategy instruction interventions on student performance (Chiu, 1998; Dignath & Büttner, 2008; Donker et al., 2014; Haller et al., 1988; Hattie et al., 1996). In these meta-analyses, the average effects of strategy instruction interventions were quite high, ranging from Hedges' g 0.40 to 0.71. Meta-analytical research on 95 interventions by Donker et al. (2014) showed that the effectiveness was moderated by several characteristics. It was found that interventions which included general metacognitive knowledge about when, why, how and which strategy to use, taught students how to plan, and addressed task value were most effective to enhance performance. Not only the “average student” profited from these forms of strategy instruction, but also students with a low socio-economic status (SES) and those with special needs benefited, the latter group seemed to profit even more. The study of Dignath, Büttner, and Langfeldt (2008) indicated that instruction consisting of a combination of metacognitive and motivational strategies yielded the highest effects. Furthermore, the effectiveness was moderated by the attributes related to the implementation and testing of the intervention. Chiu (1998) and De Boer, Donker, and van der Werf (2014) found that interventions implemented by a researcher or a research assistant were more effective than those implemented by the regular teacher or the computer, and interventions which effects were measured with standardized tests yielded lower effect sizes than when unstandardized tests were used. Dignath and Büttner (2008) found a positive effect when strategy instruction was combined with cooperative learning in secondary school (but no effect in primary school). However, De Boer et al. (2014) found slightly lower effects for cooperation. Besides that, there were indications that the effect of the intervention was related to the subject domain in which the instruction was given: interventions in the domain of writing had somewhat higher effects, and the effects in the domain of reading were somewhat lower (De Boer et al., 2014). Noteworthy, the duration of the intervention had no influence on its effectiveness (Chiu, 1998; Dignath & Büttner, 2008; De Boer et al., 2014).

1.4. Long-term effects of strategy instruction interventions

Although the findings of previous meta-studies provided input for developing effective interventions to enhance the use of learning strategies, it is still unknown what the prolonged effects of the interventions are and which type of intervention is the most effective for students' performance in the long term. Whereas all interventions included tests to measure effects right at the end of a training, very few follow-up measures were conducted. None of the previous meta-analyses on the subject (Chiu, 1998; Dignath & Büttner, 2008; Donker et al., 2014; Haller et al., 1988; Hattie et al., 1996) investigated this long-term issue.

Until now, it is assumed that findings resulting from posttests can be generalized to follow-up effects. Strategy instruction interventions that have been proven to be effective for student performance at posttest are thus assumed to also be effective in the long-term. However, we cannot know whether this is legitimate. It could be the case, for example, that some strategies need more time to develop and show an increased effectiveness in time. Furthermore, certain moderators might also affect long-term effects in a different way compared to the effects measured right after the end of the intervention program. Xin and Jitendra (1999) for example investigated long-term effects of interventions aimed at instruction in solving mathematical word problems for students with learning disabilities. About 60% of the interventions were focused on strategy training (either direct instruction in how to solve a problem, cognitive strategy use, or metacognitive strategy use). The results indicate that there was only a small decline in effect size in the period between the posttest and the follow-up test, but it was not tested whether this decline was statistically different from zero or not. The authors reported that the effect scores of interventions implemented by the researcher seemed to be slightly higher in the follow-up test than in the posttest, while the effects of interventions implemented by both teacher and researcher were lower in the follow-up. The findings of Xin and Jitendra (1999) suggest that the implementer of the intervention is likely a moderator of the follow-up effect. Xin and Jitendra (1999) furthermore checked whether the period between the posttest and the follow-up test moderated the long-term intervention effects, but this was not the case.

1.5. The current study

While several reviews and meta-analyses have shown what kind of strategy instruction is effective immediately after an intervention, a meta-analytical gap is apparent towards long-term effects of strategy instruction. As such, the current meta-analysis focuses on studies that looked at long-term effects after the intervention period had been concluded. More specifically, we focused on strategy instructions with a metacognitive component. The following research questions were formulated:

1. What are the long-term effects of metacognitive strategy instruction interventions on students' academic performance?
2. Are the long-term effects of metacognitive strategy instruction interventions moderated by the specific strategies taught?
3. Do the long-term effects of metacognitive strategy instruction interventions relate to student characteristics?
4. Are the long-term effects of metacognitive strategy instruction interventions moderated by attributes related to the implementation and testing of the intervention?

We expected positive effects for strategy instruction interventions in general and particularly for those interventions which included strategies that had proven most effective in the analysis by Donker et al. (2014). These were ‘general metacognitive knowledge’, ‘planning and prediction’, and ‘task value’. We expected that students might have become aware that these strategies enhanced their performance and, therefore, continued to use them after the experimental program had ended. This continuing use should progressively improve students’ skills in applying the learning strategies.

Regarding the other potential moderator variables, Xin and Jitendra (1999) provided insight in moderating variables for a related type of intervention studies. Furthermore, as described above, prior meta-analyses on learning strategy instruction found relationships between the instruction effect and the attributes of the implementation and testing of the intervention on short term, which may also affect the long-term effect in a similar way. In addition to the attributes that proved influential (the subject domain, the implementer of the instruction, the type of measurement instrument, the student characteristics, and cooperative learning), we included two time-related aspects, which we thought might be of relevance to the follow-up effect. These were the duration of the intervention, and the time between the posttest and the follow-up test. The duration of the intervention appeared not to have an effect on the posttest, but we wanted to be sure about the long-term effect. We expected that in interventions of longer duration learning strategy use would become more familiar. This could positively affect the students’ application of learning strategies after the end of the intervention program. The final attribute, the time elapsed between the posttest and the follow-up test, is an indicator of the maintenance of the intervention effect. If the effect declines or increases over time, this variable should be a moderator.

2. Method

Below, we describe our search procedure for the retrieval of the initial studies and the criteria for selecting the studies to be included in the meta-analysis. In addition, we explain our study coding procedure and how we analyzed the data. Fig. 1 graphically displays the search and selection procedure.

2.1. Literature search

We started by searching the internet databases *ERIC* and *PsycINFO*. The search terms we entered were ‘self-reg*’, ‘metacognit*’, ‘learning strat*’, ‘study strat*’, ‘learning skill*’, ‘study skill*’, ‘strat* use’, or ‘strat* instruction’. The search terms had to form part of the titles of the articles, or had to be mentioned as subject of the article. As advanced search options we used articles written in English and published in peer-reviewed journals. In *PsycINFO*, we selected the 12 limiters (classification codes) related to education, and in *ERIC* we selected the studies pertaining to primary or secondary school. We searched for articles published between January 2000 and January 2017. We started searching from the year 2000 as in that year Boekaerts, Pintrich and Zeidner published their handbook of self-regulation, marking a new era of research into this field. The search resulted in 8744 studies. In Refworks, we narrowed the number of studies down by additionally searching for words indicating the presence of an intervention study: ‘intervention’, ‘program’, ‘treatment’, ‘instruction’, ‘experiment’, or ‘training’. Besides this, we additionally screened the 8744 studies for one of the following terms indicating the presence of a measure of the effectiveness of an intervention a while after the end of it: ‘follow-up’, ‘delayed’, ‘maintenance’, ‘retention’, ‘long-term’, or ‘long term’, yielding a total of 4251 hits that were thoroughly screened for eligibility.

2.2. Eligibility criteria

We selected the articles in which an intervention study was described which focused on the instruction of learning strategies, including at least a metacognitive strategy (planning, monitoring, or evaluation) or metacognitive knowledge (personal or general). By including only interventions that (also) addressed students’ metacognition, we automatically excluded studies in which strategies were only taught as a ‘trick’, and the application of the strategy was a goal in itself instead of a means to enhance learning. Furthermore, the intervention had to have a minimum duration of two weeks, tested within the daily school context, and a follow-up test had to be administered to examine the long(er)-term effects of the intervention. The follow-up test had to be conducted at least three weeks after the program was finalized. We only selected articles which included the dependent variable ‘academic achievement’ (operationalized as performance in one or more school subject domains). We focused explicitly on improvement of academic achievement, and did not include other types of available measures. We excluded correlation studies which only examined the relationship between strategy use and student achievement. In these studies, strategy instruction was not implemented through training, meaning that we would be unable to analyze the possible causal relationship between learning strategy instruction interventions and student achievement. With respect to the subject domain, we limited our study to the core academic subjects. Subjects such as music, arts, and physical education were excluded, as we were particularly interested in the effects of interventions focused on academic achievement in the more cognitive domains. The research sample had to consist of primary school or secondary school students, from grade one up to and including the twelfth grade, following most European and the American school systems. Students in higher education were excluded, as we wanted to know if it is possible to teach students metacognitive strategies with sustaining effects on student performance already in the mandatory school years. Furthermore, we used the following methodological criteria in order to select studies of sufficient quality:

- The research had to include a control group. If there was no control group, it would be unclear whether the results of the experimental group were caused by the intervention or by normal developments. It was allowed that the control group also had

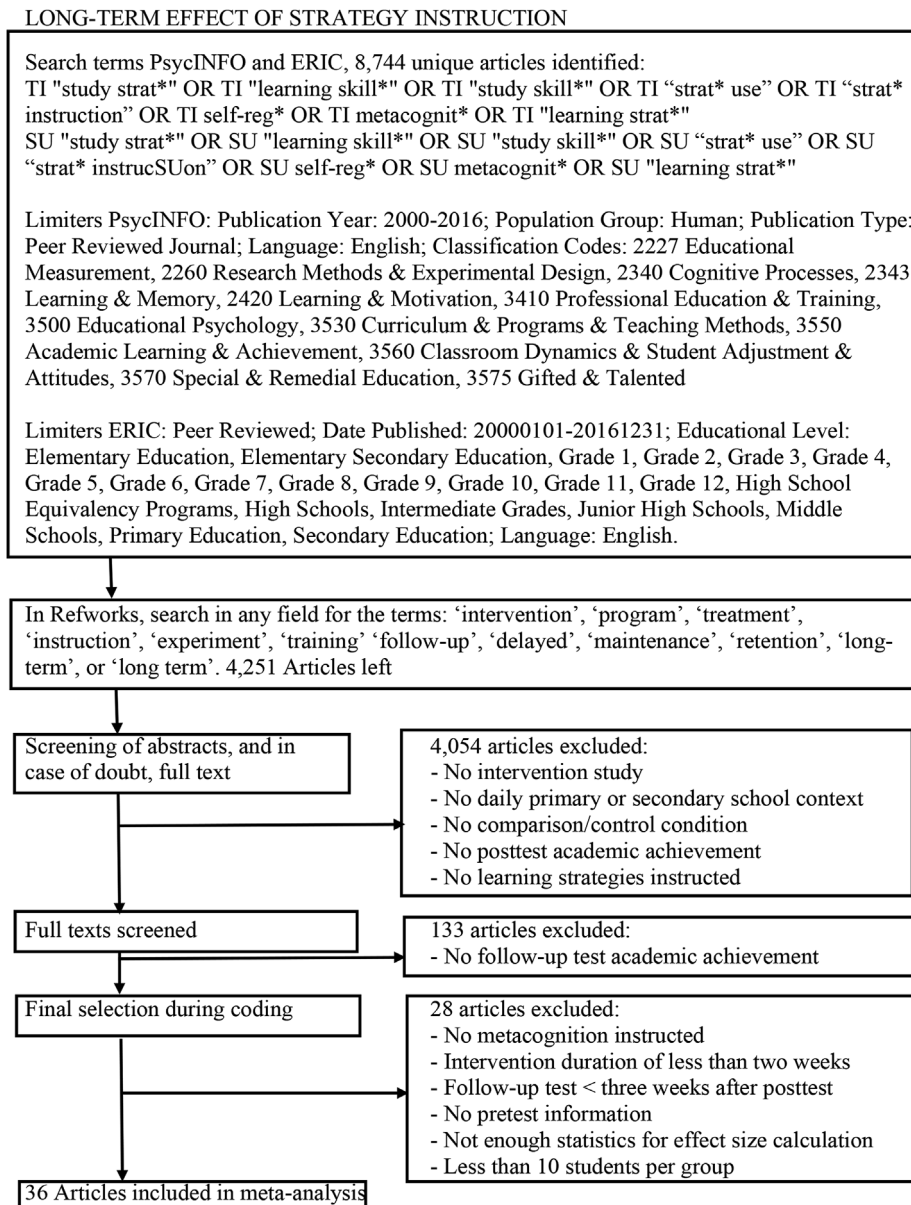


Fig. 1. Literature search and selection procedure.

some type of intervention, for example if there was no 'business as usual' comparison group available. In these cases, we only coded the learning strategies that the experimental group were taught additionally compared with the control group.

- The research had to provide pretest, posttest, and follow-up test measures. Studies which did not provide pretest scores were only included if it was indicated that there were no initial differences between the control and the experimental group.
- The study samples had to include at least ten students per group in order to ensure that the effect size Cohen's *d* would be approximately normally distributed (Hedges & Olkin, 1985). Studies with fewer than ten students per group were, therefore, excluded from our meta-analysis.

The literature search yielded 36 articles that met our eligibility criteria. Two articles examined the same intervention and sample, but focused on a different achievement test. These were the articles of Ben-David and Zohar (2009) and Zohar and Ben-David (2008). We counted these articles as being one intervention. Several of the 36 articles, however, described more than one intervention, resulting in a total of 48 interventions for our meta-analysis. Although we focused on interventions implemented in primary school and secondary school, the selection included only interventions aimed at students in grades one to eight, so the higher secondary school grades were not represented in the study.

2.3. Coding

The first two authors coded all articles; inter-coder reliability was high, with a 96% agreement. Doubts about the eligibility or coding of a study were discussed until agreement was reached. Table 1 presents an overview of the studies and their characteristics. The variables relevant to the current meta-analysis were:

Measurement instrument. We distinguished two types of instruments used to assess the effects of the strategy instruction: standardized tests and unstandardized tests.

Subject domain. We coded the subject domain in which the training was offered. We distinguished (comprehensive) reading, writing, mathematics, science, and a category labelled ‘other’.

Duration of the intervention, coded in weeks.

Time between posttest and follow-up test, coded in weeks.

Implementer of the intervention. As we were particularly interested in the difference between the teachers as implementers and others, we distinguished between the regular teacher as implementer and ‘other implementers’. These others could be the researcher, an assistant researcher, or the computer.

Cooperation. In terms of whether the training was focused on cooperative or on individual learning, we distinguished two categories: ‘cooperation in intervention group but not in control group’ and ‘both groups do or do not cooperate’.

Student characteristics. We distinguished ‘regular students’ (when interventions were aimed at students who had none of the characteristics mentioned below, or when no specific information was provided, in which case we assumed that the students were representative of the majority of the population which was the focus of the research), ‘low SES students’, ‘high SES students’, ‘students with special needs’, and ‘gifted students’. However, our search yielded no strategy instruction interventions aimed at high SES or gifted students, therefore these categories were not applicable after coding all articles.

Learning strategies. In line with our theoretical framework, we distinguished three types of learning strategies, as well as meta-cognitive knowledge and motivational aspects. In total, we coded for fourteen categories, defined as follows:

Metacognitive knowledge.

1. *Personal metacognitive knowledge.* The student's knowledge of his/her own learning. This knowledge relates to one's personal strengths and weaknesses, and how these can compensate one another. It particularly concerns information on how ‘the student him/herself’ learns best.
2. *General metacognitive knowledge.* Knowledge of learning and cognition in general, including knowledge of how, when, and why to use learning strategies.

Cognitive strategies. The following types of cognitive strategies were coded:

3. *Rehearsal.* Repeating and re-reading words and text passages in order to remember the contents and be able to use them.
4. *Elaboration.* Actively making connections between new and already known information and structuring the material in order to facilitate the storage of this knowledge in the long-term memory.
5. *Organization.* Reducing information to the relevant issues to enhance one's comprehension. Examples: categorizing information, structuring a text, transforming text into a graph.

Metacognitive strategies. We distinguished three types of metacognitive strategies related to the phases of the learning process:

6. *Strategies for planning and prediction.* An explicit focus on planning and the use of time, based on which the students can predict how they are going to perform and what they will need in order to perform well. Examples: making a plan, starting with the most important aspect, and determining how much time one will need to spend on the task.
7. *Strategies for monitoring and control.* Monitoring the learning process by checking if one is still ‘on the right track’ and adjusting one's learning approach if required. Examples: formulating questions to check one's understanding, checking information.
8. *Strategies for evaluation and reflection.* After completing a task, reconsidering either the process or the product. Examples: checking answers before handing in an assignment, comparing the outcome to the goal.

Management strategies. Three categories of management were distinguished:

9. *Management of the self, or effort management.* This concept is related to motivation. It reflects the commitment to reaching one's study goals even when there are problems or distractions. Examples are goal-directed behavior and perseverance despite difficulties.
10. *Management of the environment.* Looking for possibilities in the environment to create the best circumstances for learning, for instance, finding a quiet place to study, but also using dictionaries and going to the library – or the internet – to look for information.
11. *Management of others.* Help-seeking.

Motivational aspects. Regarding motivation, we distinguished the following three categories:

12. *Self-efficacy*. A student's belief in his or her ability to successfully complete a task. Includes judgments about one's ability to accomplish a task as well as confidence in one's skills to perform the task.
13. *Task value*. Belief in the relevance and importance of a task.
14. *Goal orientation*. The degree to which the student perceives him/herself to be performing a task for reasons such as seeking a challenge, curiosity, wanting to master a skill (intrinsic), or obtaining high grades, getting rewards, receiving a good performance evaluation from others, or competition (extrinsic).

2.4. Meta-analysis

2.4.1. Combining the effects and moderator analysis

To perform the meta-analysis we used the statistical packages Comprehensive Meta-Analysis (CMA) version two, developed by Biostat (see: www.meta-analysis.com), and Hierarchical Linear Modeling (HLM) version six, developed by Raudenbush, Bryk, and Congdon (2004). CMA was used to compute the effect sizes (Hedges' g) and variances of the individual interventions based on the statistical information provided in the primary studies. In several studies, the intervention effect was estimated using more than one test. We included all these measures in our meta-analysis, and let CMA calculate the mean effect. The statistical package CMA was also used to calculate average weighted effect sizes for the summary effect of all studies, analyze publication bias, and perform meta-ANOVA: an analysis of variance for meta-analytical data to examine whether the intervention effects were moderated. Because the interventions included in the meta-analysis differed in many respects, we used a random effects model to estimate the weighted average effect size. We used a mixed effects model for the moderator analyses. HLM was applied to perform meta-regression analysis with multiple predictors.

2.4.2. Dependency of the data

In several studies, multiple interventions were tested and compared using the same control group. This causes statistical dependency in the data (Lipsey & Wilson, 2001), and if we failed to correct for this, it would result in too much weight being attached to the control group. We allowed for this by dividing the number of students in the control group by the number of interventions with which the control was compared. As a result, the variance increased, and the weight – which is the inverse of the variance – decreased.

2.4.3. Extreme effect sizes

Following Lipsey (2009) we recoded extreme effect sizes of individual studies by means of the Tukey method. Outliers more than 1.5 times the interquartile range beyond the 25th and 75th percentiles were recoded to the respective inner fence values. For the post-follow-up effect, this meant that three extremely negative effects were recoded (these were from the studies of Ariès, Groot, & Maassen van den Brink, 2015; Fidalgo, Torrance, & Garcia, 2008; Jitendra, Hoppes, & Xin, 2000, with effect sizes of Hedges' $g = -1.24, -1.57$ and -1.15 respectively, and recoded into -0.625). For the pre-post effect and for the pre-follow-up effect, respectively five and four interventions with extreme effects (positive and negative) were recoded.

2.4.4. Type 1 error correction

We performed multiple moderator analyses, which increases the likelihood of Type 1 errors: incorrectly rejecting the null hypothesis that there is no effect. In his dissertation on the issue of multiple statistical significance tests in meta-analyses, Polanin (2013) suggested applying the 'false discovery rate' (FDR) method of Benjamini and Hochberg (1995) within each group of tests ('timeline of significance testing'). This method adjusts for the increased chance of Type 1 errors, while balancing the chances of incorrectly rejecting the null hypothesis (Type 1 error) and falsely accepting the null hypothesis (Type 2 error). In accordance with the FDR procedure, we ordered the p -values within each group of tests (summary effect, categorical moderator analyses, continuous moderator analyses, multiple meta-regression) from low to high, starting with the largest value, and determined the largest p -value for which

$$p_i \leq \frac{i}{m} * \alpha$$

where i is the ordered p -value, m the number of significance tests and α the chosen level of control. We chose $\alpha = 0.05$. We then rejected the null hypotheses of p_i and smaller.

3. Results

3.1. Average long-term effect of strategy instruction on student academic performance

The meta-analysis included 48 strategy instruction interventions in which the long-term effects on student academic performance were measured. The follow-up tests, by which the long-term effects were measured, were often administered at the moment at which about a similar number of weeks had been passed after completion of the intervention as the period of intervention itself had lasted. So, if an intervention took five weeks, the follow-up test was administered about five weeks after it had ended. On average, the follow-up test was conducted after 21.6 weeks ($SD = 23.5$). The shortest period was three weeks and the longest period 108 weeks.

The average effect sizes (Hedges' g) of strategy instruction measured right after the end of the interventions and at the time of the follow-up test were 0.50 ($SE = 0.07$) and 0.63 ($SE = 0.07$), respectively. These are moderately large effects (following the interpretation of Cohen, 1988).

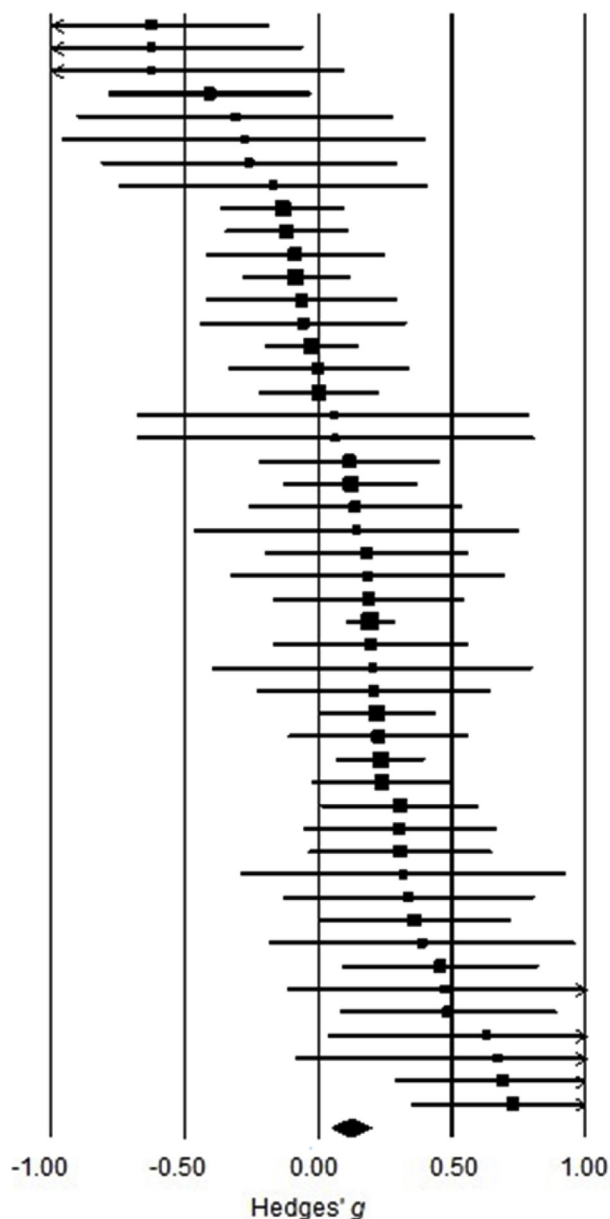


Fig. 2. Forest Plot of the Average Follow-Up Effect and the 95% Confidence Interval of Each of the Strategy Instruction Interventions (Represented by a Square; Interventions with Larger Sample Sizes are Represented by a Bigger Square) Plus the Summary Effect (Represented by a Diamond).

To examine the differences between student performance at the posttest and the follow-up test more thoroughly, we calculated the effect sizes of these differences for all interventions. Based on the measures at the posttest and the follow-up test, the average (weighted) effect size of this difference was Hedges' $g = 0.12$ ($SE = 0.04$), with a 95% confidence interval of Hedges' $g = 0.05$ to 0.19 . This is a very small but significant effect ($p = 0.001$). So we see that strategy instruction interventions did have a sustained effect on student performance, and the relative gain in student performance in the experimental groups compared with that in the control groups had increased even slightly from posttest to follow-up test. As we were especially interested in what happens with the student performance scores between the posttest and the follow-up test, we used the measure based on the posttest and the follow-up test scores in the remainder of the analyses. When we use the term follow-up effect, we are referring to this post-follow-up difference measure. The homogeneity statistic of the follow-up effect indicated that the variation in effects is statistically significant ($Q = 99.06$; $df = 47$; $p < 0.001$), which shows that the strategy instructions do not share the same true effect size. The variance of the true effect sizes is estimated at $T^2 = 0.028$. The variance reflects a moderate proportion of real differences in effect size, as indicated by the I^2 of 52.55. Fig. 2 depicts the average follow-up effect and its 95% confidence interval for each of the 48 strategy instruction interventions. The interventions are sorted from low to high effect size.

To reveal the degree of possible publication bias in the meta-analysis, we used Duval and Tweedie's *Trim and Fill* method (Borenstein, Hedges, Higgins, & Rothstein, 2009; Peters, Sutton, Jones, Abrams, & Rushton, 2007). This method is based on the idea that publication bias leads to an asymmetric funnel plot, which means that the effect sizes of the primary studies in a meta-analysis are not distributed evenly around the mean effect. The Duval and Tweedie's method explores if the symmetry can be optimized by imputing (filling in) trimmed values of the most extreme effect sizes, but with opposite effect direction. If the symmetry can be optimized by imputed values, this indicates that there were studies missing in the meta-analysis, suggesting the presence of publication bias. We used a random effects model to determine whether there were interventions missing in the meta-analysis. We have looked to the left side as well as to the right side of the mean and found three interventions missing with a negative effect size. The estimated effect size without publication bias would have been slightly lower with a Hedges' $g = 0.09$ ($SE = 0.04$), but still a statistically significant effect.

3.2. Moderation of the long-term effect by intervention characteristics

The homogeneity analysis indicated that the strategy instruction interventions do not share the same true effect size. We examined whether the specific learning strategies taught, the student characteristics of the sample, and attributes of the implementation and testing of the intervention explained the variance in effect sizes. We started by analyzing the attributes of the implementation and testing.

3.2.1. Measurement instrument

We first investigated the influence of the type of instrument used for measuring student performance. For exactly half of the 48 instructions, the effects were estimated with standardized achievement tests. The other half used an unstandardized test. It appears that the effects measured using standardized tests increase between the posttest and the follow-up test (Hedges' $g = 0.17$; $SE = 0.05$; $p < 0.01$), whereas the effects measured using unstandardized tests remain stable (Hedges' $g = 0.07$; $SE = 0.06$; $p = 0.27$). However, the difference in effect between both types of measurement instruments is not statistically significant (Q -between = 1.85; $df = 1$; $p = 0.17$).

3.2.2. Subject domain

Table 2 lists the average follow-up effects per subject domain. The results show small differences between the subject domains, but a *meta*-ANOVA indicated that these differences were not statistically significant. Post hoc analyses revealed that instructions implemented during science lessons had significantly lower effects than instructions implemented in math ($p = 0.02$). There were no further differences between two subject domains.

3.2.3. Time aspects

Here we looked at the duration of the intervention and the time between the posttest and the follow-up test. The average duration of the instructions was 19.8 weeks ($SD = 21.9$), and the average time between the posttest and the follow-up was 21.6 weeks ($SD = 23.5$), as also reported above. Using meta-regression, we tested whether the duration of the intervention, and the time between the end of the strategy instruction and the follow-up test influenced the follow-up effects. None of the predictors had a statistically significant effect, however. The regression coefficients B (and standard errors) were 0.00 (0.00) for both moderators.

3.2.4. Implementer of the intervention

The average effect of the 33 strategy instruction interventions implemented by the teacher was Hedges' $g = 0.12$ ($SE = 0.04$; $p < 0.01$), and that of the 15 instructions implemented by others was almost the same with 0.11 ($SE = 0.07$; $p = 0.13$). *Meta*-ANOVA indicated no statistically significant difference (Q -between = 0.03; $df = 1$; $p = 0.87$). The group 'other implementer' consisted of three interventions implemented by the researcher, 11 by the research assistant, and one by means of the computer.

Table 2
Influence of the subject domain on the follow-up effect.

Variable	n Interventions	Hedges' g (SE)
Reading	24	0.12 (0.04)*
Writing	6	0.05 (0.12)
Math	8	0.22 (0.08)*
Science	5	-0.03 (0.07)
Other	5	0.12 (0.29)

Meta-ANOVA: Q -between = 6.11; $df = 4$; $p = 0.19$.

Notes. The category 'other' includes two instructions focused on social sciences, two instructions on all subjects, and one instruction on listening comprehension.

* $p < 0.05$.

Table 3
Influence of each learning strategy on the follow-up effect: Meta-regression results.

Learning strategy	<i>n</i> Interventions	Strategy included Hedges'g (SE)	Strategy not included Hedges'g (SE)	<i>Q</i> -between; <i>df</i> ; <i>p</i>
Metacogn. knowl. personal	4	−0.10 (0.12)	0.14 (0.04)*	3.53; 1; 0.06
Metacogn. knowl. general	28	0.10 (0.05)*	0.15 (0.06)*	0.41; 1; 0.52
Rehearsal	3	−0.23 (0.12)	0.15 (0.04)*	8.746; 1; 0.003*
Elaboration	30	0.12 (0.05)*	0.13 (0.07)	0.03; 1; 0.87
Organization	17	0.11 (0.06)	0.13 (0.05)*	0.06; 1; 0.82
Planning and prediction	32	0.10 (0.04)*	0.16 (0.08)*	0.42; 1; 0.52
Monitoring and control	38	0.10 (0.04)*	0.18 (0.09)	0.56; 1; 0.45
Evaluation and reflection	36	0.10 (0.04)*	0.20 (0.08)*	1.26; 1; 0.26
Effort management	2	0.33 (0.16)*	0.11 (0.04)*	1.80; 1; 0.18
Manage environment	0	–	–	–
Manage others	7	0.20 (0.07)*	0.11 (0.04)*	1.23; 1; 0.27
Self-efficacy	4	0.07 (0.17)	0.12 (0.04)*	0.11; 1; 0.74
Task value	8	0.14 (0.10)	0.12 (0.04)*	0.03; 1; 0.87
Goal orientation	0	–	–	–

Note. * $p < 0.05$.

3.2.5. Cooperative learning

We did not find an influence of cooperative learning on the follow-up effects of interventions. The nine interventions in which the intervention group cooperated but the control group did not, had an average effect of 0.13 ($SE = 0.10$; $p = 0.20$). The 39 interventions in which both groups either did or did not cooperate had an average effect of 0.12 ($SE = 0.04$; $p < 0.01$). The meta-ANOVA results were Q -between = 0.014; $df = 1$; $p = 0.90$.

3.2.6. Student characteristics

Our meta-analysis included 36 interventions for regular students, four for low SES students, and eight for special needs students. The average follow-up effects for these groups were 0.10 ($SE = 0.04$; $p = 0.02$), 0.35 ($SE = 0.11$; $p < 0.01$) and 0.04 ($SE = 0.13$; $p = 0.75$), respectively. With a Q -between of 5.14 ($df = 2$; $p = 0.08$), the between-groups differences were not statistically significant. Post hoc analyses showed a significant difference between low SES students and average students ($p = 0.03$).

3.2.7. Learning strategies

Finally, we present the results of the analysis of the influence of the learning strategies taught during the interventions on the follow-up effects as regards student performance. For each of the learning strategies coded, we performed a separate meta-ANOVA, with in each analysis only the strategy of interest as moderator (coded as a dummy variable with code 0 'not in intervention' and code 1 'instructed in intervention'). Table 3 presents the results and shows that only for the cognitive strategy rehearsal the Q -between value indicated a statistically significant difference. Instructions with the strategy rehearsal had lower follow-up effects than instructions without this strategy.

3.2.8. Simultaneous analysis of multiple moderators

Finally, we performed two meta-regressions with multiple moderators. As the instructions were characterized by many different aspects, a meta-regression with multiple moderators is the best way to analyze the effects of these aspects in conjunction. The number of interventions included in the meta-analysis restricted the number of moderators that could be included in the model. Therefore, we had to make a selection of moderators that seemed most relevant.

The first model included the moderators for which we had found an effect. We found effects for low SES students versus average students, for science versus math, and for rehearsal. We therefore included the subject domain, the student characteristics and the learning strategy rehearsal in the same meta-regression to explore how these aspects together contribute to the variance in effect sizes. Model 1 in Table 4 shows the results of the meta-regression. Compared with the empty model, this model explained 57.98% of the variance in follow-up effects. We found statistically significant differences for the student characteristics, for rehearsal and for the subject domain. Follow-up effects were higher for low SES students than for average or special needs students. Furthermore, interventions including rehearsal had lower follow-up effects than interventions without this strategy type. Besides that, we found that interventions in writing had lower effects than interventions in math and the category other subjects.

The second meta-regression was an attempt to analyze all the different learning strategies in conjunction with the attributes for which we found statistically significant effects. To keep some balance in the number of interventions included in the meta-analysis and the number of moderators included in the meta-regression, we used the categories of strategy types instead of the individual strategies. Table 4, Model 2, presents the results. We found statistically significant effects for metacognitive knowledge and the student characteristics. Including metacognitive knowledge in the intervention yielded a lower effect. We found higher effects for low SES students compared with average students or special needs students.

Table 4
Meta-regression of multiple moderators on the follow-up effect.

Variable	<i>n</i> Interventions	Model 1 Hedges' <i>g</i> (SE)	Model 2 Hedges' <i>g</i> (SE)
Intercept		0.18 (0.11)	0.53 (0.27)
Reading vs. writing	24 vs. 5	0.20 (0.11)	0.10 (0.13)
Math vs. writing	6 vs. 5	0.26 (0.11)*	0.20 (0.13)
Science vs. writing	8 vs. 5	0.24 (0.17)	0.12 (0.19)
Other subjects vs. writing	5 vs. 5	0.35 (0.17)*	0.01 (0.13)
Average students vs. low SES	4 vs. 36	−0.25 (0.11)*	−0.45 (0.13)*
Special needs vs. low SES	8 vs. 36	−0.35 (0.17)*	−0.53 (0.19)*
Cognitive strategy rehearsal	3	−0.50 (0.19)*	
Metacognitive knowledge	32		−0.22 (0.09)*
Cognitive strategies	31		0.03 (0.08)
Metacognitive strategies	46		0.05 (0.18)
Management strategies	9		0.18 (0.12)
Motivational aspects	11		−0.15 (0.10)
Variance		0.01426	0.01784

Note. * $p < 0.05$.

3.2.9. Applying the correction for type 1 errors

We performed many significance tests and found several effects for which the p -value was below 0.05. However, after applying the False Discovery Rate method, only a few effects were likely not false discoveries and reflected real differences. This applied to the overall summary effects, for the *meta*-ANOVA analysis with the cognitive strategy rehearsal as the single moderator, and furthermore to the finding from the final meta-regression that follow-up effects are higher for low SES students compared with average students.

3.2.10. Sensitivity analysis

Three interventions with extremely negative effect sizes for the post-follow-up difference were recoded. We also performed the above analyses without recoded values for these extreme effects to investigate how our recoding affected the findings. The conclusions were, however, comparable with the findings described above, although some effects were not statistically significant anymore. In first instance, we found significant effects for the overall summary effects from pretest to posttest and from pretest to follow-up test. However, with a p -value of 0.056, the effect from posttest to follow-up test was just not statistically significant. With respect to the moderators, effects were found for the student characteristics and the cognitive strategy rehearsal. The meta-regression with the moderators for which we found effects again showed an effect for rehearsal. The meta-regression with the attributes for which we found significant differences and the various strategy types showed that effects were higher for low SES students compared with other students, and that instructions with metacognitive knowledge had lower effects than instructions without. After applying the type 1 error correction, only the summary effects and the meta-regression effect for rehearsal were classified as true effects and not false discoveries. The finding from the final meta-regression with respect to the higher effects for low SES students compared with average students, just did not pass the FDR-procedure with a 0.002 too high p -value.

The results of the analyses with and without the recoding of extreme effect sizes point in the same direction. The analyses without the adjustment of extreme effect sizes, however, likely suffer from those extremes. Extreme effects distort the findings in an unrepresentative way, by their disproportionate influence on the (group) means and variances (Lipsey & Wilson, 2001). We therefore think it is wise to base our conclusions on the analyses with the adjustment for extreme effect sizes, and we feel these are justified by the sensitivity analysis because the findings are not very different.

4. Discussion

Several meta-analyses have been conducted to examine the effects of learning strategy instruction interventions on student academic performance. In none of these studies, however, was it investigated whether these interventions have long-term positive effects. In the current meta-analysis we addressed this question, and also investigated whether particular characteristics of the interventions moderate the long-term effects of strategy instruction on student performance. All studies included in the meta-analysis at least addressed a metacognitive strategy or metacognitive knowledge.

4.1. Long-term effect

The 48 interventions included in the meta-analysis had an average effect size of Hedges' $g = 0.50$ ($SE = 0.07$) on student academic performance at posttest. The average long-term effect size measured in the follow-up tests appeared to be Hedges' $g = 0.63$ ($SE = 0.07$), which is a moderately large effect size. The small difference in effect between the posttest and the follow-up test, Hedges' $g = 0.12$ ($SE = 0.04$), was statistically significant. This result indicates that the effect of strategy instruction, with at least a metacognitive component, is mainly achieved during the period of instruction, but that even after the instruction has ended, the effectiveness slightly increases. It seems as if students' learning skills further developed after the end of the instruction. The only prior

meta-study which lends itself to some comparison of long-term learning strategy instruction effects is that of [Xin and Jitendra \(1999\)](#). Their results suggest that the effects of instruction programs in mathematical word problem-solving for students with learning problems declined slightly over time. However, this decline was small and was not tested for significance. The finding of our current meta-analysis that the effects of learning strategy instruction interventions are maintained and slightly gain over time, at least during the timespan measured, is a promising result. It shows that the interventions reach their goal in that student performance is sustainably enhanced.

4.2. Moderators of the long-term effect

Next, we examined whether intervention characteristics moderated the long-term effects. To perform the moderator analyses, we computed the long-term effects using the scores measured during the posttest and the follow-up test. The difference in scores between the two measurement moments indicated the follow-up effect. We analyzed the moderation effects of the following characteristics: the type of measurement instrument used to evaluate the intervention effects, the subject domain in which the strategy instruction was given, the duration of the intervention, the time between the posttest and the follow-up test, the implementer of the intervention (teacher or other), whether student cooperation was stimulated, the characteristics of the students, and the learning strategies taught. The findings showed that there were two moderators of the long-term effect.

4.2.1. Learning strategies as moderators of the long-term effect

The first moderator of the long-term instruction effect was the teaching of the cognitive strategy rehearsal. Instructions in which this strategy was taught had lower long-term effects than instructions without this strategy. Rehearsal is a cognitive strategy that is only useful for shallow processing of the learning material. Other strategies, like elaboration and organization, focus on deeper processing and thus a deeper understanding of the learning material ([Mayer, 2008](#)). It might be for this reason that rehearsal is found to be a less effective learning strategy. Besides that, rehearsal is a strategy that is probably most often already part of a student's strategy repertoire, so extra instruction does not add much.

We found no moderator effects for the other learning strategies on the long-term instruction effect. [Donker et al. \(2014\)](#) reported that interventions including general metacognitive knowledge, planning and task value were more effective than others in enhancing student performance. We expected that the effects of interventions which included one or more of these strategies would increase after completion of the experimental trajectory, because the students would continue to use these strategies, thereby improving their skills. Our findings, however, do not confirm this hypothesis. We did find an overall small increase of the effect between posttest and follow-up measure, but not an additional effect for specifically these strategies. [Donker et al. \(2014\)](#) suggested that some learning strategies might require longer periods of time before students could use them properly. This was suggested because it was found that interventions containing particular strategies had on average lower effects than interventions which did not include these approaches. This was the case for goal orientation, and in reading comprehension interventions for elaboration and the management strategy 'peers/others'. The current findings, however, showed that the effectiveness of interventions including elaboration and the management of others did not increase between the posttest and the follow-up test. Unfortunately, we had no interventions including goal orientation. Therefore, we were unable to test our hypothesis with respect to this strategy.

4.2.2. Student characteristics as moderator of the long-term effect

The second moderator of the long-term instruction effect was the student characteristics. Long-term effects of strategy instructions were higher for low SES students than for other students. A bulk of research on the relationship between SES and academic achievement, summarized by [Dietrichson, Bøg, Filges and Klint Jørgensen \(2017\)](#), has shown that low SES students have lower achievement levels than average and high SES students. This lower achievement level of low SES students does not so much relate to nature (intelligence), but to nurture (environment). [Dent and Koenka \(2016\)](#) showed a positive relationship between metacognitive and cognitive strategy use and students' academic achievement levels. Based on these findings, we hypothesize that low SES students were less equipped with learning strategies at the start of the instructions than middle and higher class students. So, if low SES students are less equipped with learning strategies by their home environment (nurture) and this is one of the reasons why they have lower achievement levels than other students with higher SES and comparable intelligence, it means that there is 'more to win' for the low SES group. Instructing learning strategies to low SES students therefore possibly resulted in a greater gain in learning skills and, as a consequence, in higher academic performance. However, if strategy instruction has indeed a greater impact for low SES students, we expected that the one prior meta-analysis that also focused on this aspect, of [Donker et al. \(2014\)](#) would have also found this effect already at posttest. This was not the case, though. We additionally examined the pre-posttest differences with respect to the student characteristics for the current study, and did find a higher effect for low SES students than for average students, although the difference was not statistically significant (Hedges' g of 0.68 and 0.44, respectively). It therefore remains unresolved why we found a higher long-term effect for low SES students.

4.2.3. Attributes of the implementation and testing of the intervention as moderators of the long-term effect

None of the attributes related to the implementation and testing of the intervention was found to have a moderator effect. [Xin and Jitendra \(1999\)](#) also tested for moderation effects and found, just as in the current study, no connection between the time period between the posttest and the follow-up test and the long-term effect. Similar to our findings, their results further suggested that the duration of the intervention was not a statistically significant moderator of the follow-up effects. Only interventions implemented by the researcher seemed to have positive long-term effects. Unfortunately, however, Xin and Jitendra did not include any interventions

in their meta-analysis that were implemented by the teacher only and in which follow-up tests were used. Therefore, their findings with respect to the implementer of the intervention cannot be compared with our results.

4.3. Limitations

It may have confused our findings that the effect sizes of the interventions were measured using a large spectrum of different tests. These tests were of course not calibrated, so the students' performance was measured in several different ways. In addition, we noticed that there were also sometimes large differences between the effect sizes within the individual interventions. This made it more difficult to compare the effect sizes and analyze the moderator impact.

Furthermore, as they always are, the selection criteria were in some respects arbitrarily. This applied in particular to the minimum duration of the instruction and the time between the posttest and follow-up test, which we set at two and three weeks respectively. It implied that we rejected studies that had five daily lessons during one week, and included a study with two lessons spread over two weeks. Our choices were led by the assumption that, for long-term effects, students needed at least some time to process the new information, and there had to be at least a few weeks between the posttest and follow-up test. However, in order to include a sufficient number of studies in the meta-analysis, we could not be too stringent. The results indicated that the duration and time elapsed between the posttest and follow-up test had no influence on the long-term effect, which we think makes our choices acceptable.

One may argue whether the effects we found were actually evoked by the metacognitive strategy instruction or by something else, as we did not analyze the direct effects on metacognitive skillfulness. However, we purposefully did not focus on metacognitive skillfulness as outcome measure. Metacognitive skillfulness is, in our view, not an end in itself, but a means to enhance learning and thus students' performance. Our interest was therefore in the effects on students' performance and not on metacognitive skillfulness. Besides that, metacognitive skillfulness is not a treatment independent outcome measure, as only the experimental group received the metacognitive strategy instruction and the control group did not. Comparing the experimental group with the control group on metacognitive skillfulness would yield a biased effect. In line with the Best Evidence Organization, we thought it not to be appropriate to include treatment inherent measures in the meta-analysis (<http://www.bestevidence.org/aboutbee.htm>). We think, though, that we have overcome the issue whether the effects were evoked by the metacognitive strategy instruction or not, by making use of control groups with which we compared the experimental groups. The differences between the experimental and control groups pertained the instructed strategies, and not the amount of learning or instruction time, and neither the subject that was taught. As such, we think it is appropriate to assume that it was the strategy instruction that caused differences in student performance between the experimental and control groups.

A last aspect that complicated the moderator analysis was that the meta-analysis included a restricted number of interventions that met our criteria. This limited our statistical options. For example, in a meta-analysis, approximately one variable can be added per ten interventions (Borenstein et al., 2009). With 48 interventions, we could only include approximately five variables in the meta-regression equation at any one time. Therefore, we do not know whether our results were influenced by the exclusion of a particular variable. The restricted number of studies also sometimes resulted in a comparison between two or more groups of studies, distinguished by the scores on a moderator variable, with only a small number of studies in one or more of the distinguished groups. This makes it more difficult to find effects, but it also makes the analyses more susceptible to incidentally large effects of individual studies; although we partially reduced this last side effect by recoding extreme effect sizes. Our findings with respect to the moderating effects of the cognitive strategy rehearsal and SES were indeed based on only three studies with the strategy rehearsal, and only four studies focused on low SES students. The individual effects of these studies were however in congruence with our conclusion. All studies with rehearsal had lower than average long-term effects, and all studies focused on low SES students had effects that were higher than the average.

4.4. Scientific and practical contributions

Meta-analysis is a valuable method for summarizing the findings of primary studies. As such, the current study has contributed to the body of knowledge about learning strategy instruction. Despite the limitations of meta-analytical research, an analysis of several primary publications yields more reliable conclusions than that of a single primary study. In meta-analysis, measurement and implementation errors of the primary studies are averaged out, which leads to more balanced results.

The current meta-analysis has clearly shown that metacognitive strategy instruction has sustained effects on student academic performance; the gain in student performance as a result of a strategy instruction intervention is maintained after the trajectory has ended, and even slightly increased at long-term (except when it included the cognitive strategy rehearsal). It shows that students maintain their newly acquired learning skills and it indicates that they can be prepared for self-regulated learning after the official school period. This will help students to become lifelong learners, which is necessary in the current complex society, which continuously requires new knowledge and skills. It is therefore useful to implement metacognitive strategy instruction in the school program. The sustainability of the instruction effect indicates that teachers do not have to address strategy instruction in each lesson. However, Hattie et al. (1996) found that students have difficulties transferring learning skills to new tasks. Therefore, we do recommend teachers to provide strategy instruction for each new learning task. Furthermore, the current meta-analysis indicated that most learning strategies have similar sustainable effects, as most strategies had no moderating influence on the long-term instruction effect. It is likely that each specific learning task has its own palette of effective strategies. We therefore recommend teachers to instruct students a wide range of learning strategies, so that students are well equipped for all kinds of learning tasks.

The study showed that in particular low SES students benefit from strategy instruction in the long term, as their performance levels increased more than that of other students. As such, strategy instruction seems a valuable instrument in reducing the well-known achievement gap between students with a low and a higher socio-economic status, although more research is needed to make firm conclusions.

4.5. Future research recommendations

The aim of teaching students learning strategies is to enable them to acquire new knowledge and skills by themselves; not only during an intervention period, but also in the future. The aim is therefore long-term. Although there is an overwhelming amount of learning strategy instructions tested for their effectiveness, we found only a relatively small amount of studies that focused on longer-term effects. As these long-term effects are so important, we would recommend future researchers of learning strategy instructions to also examine these effects. A secondary benefit of these long-term measures in future primary studies will be that there become more studies available for a meta-analysis of possible moderators of the long-term effect.

Furthermore, our research indicated that low SES students benefitted more from strategy instruction. More research is needed, however, to find out why this group of students profited extra. Do low SES students indeed start with less learning skills than other students? Instructions tested on low as well as on higher SES students, in which pretest, posttest and follow-up measures are provided for both groups will give more insight into this question.

Lastly, to our surprise we barely found studies based on computerized interventions (more specifically: only in one intervention the instructor was coded as “computer”, instead of the teacher or another person such as the researcher). Over the last years, computerized instruction has taken a flight in research on the use of computers as metacognitive tools for enhancing learning (e.g., Azevedo, 2005). For example, the meta-analysis of Lan, Lo, and Hsu (2014), which investigated the effects of metacognitive instruction on students' reading comprehension in computerized reading contexts, found potential for such instruction and the study of Lenhard, Baier, Endlich, Schneider, and Hoffmann (2013) showed that of two strategy training programs, the one delivered computer based, was even more effective. This type of interventions did not surface in our literature search, possibly due to our search terms, yet is an important, related, topic for future research. Given the fact that computerized instruction becomes more mainstream in education and given the rich possibilities for more personalized instruction in such environments, this could play an important role in fostering students' metacognition and use of learning strategies.

References

- Studies included in the meta-analysis are preceded by an asterisk.
- Alexander, P. A., Graham, S., & Harris, K. R. (1998). A perspective on strategy research: Progress and prospects. *Educational Psychology Review*, 10, 129–154. <http://dx.doi.org/10.1023/A:1022185502996>.
- (*)Alfassi, M. (2002). Communities of learners and thinkers: The effects of fostering writing competence in a metacognitive environment. *Journal of Cognitive Education and Psychology*, 2, 119–135. <http://dx.doi.org/10.1891/194589502787383335>.
- (*)Antoniou, F., & Souvignier, E. (2007). Strategy instruction in reading comprehension: An intervention study for students with learning disabilities. *Learning Disabilities: A Contemporary Journal*, 5, 41–57.
- (*)Ariès, R. J., Groot, W., & Maassen van den Brink, H. (2015). Improving reasoning skills in secondary history education by working memory training. *British Educational Research Journal*, 41, 210–228. <http://dx.doi.org/10.1002/berj.3142>.
- Azevedo, R. (2005). Computer environments as metacognitive tools for enhancing learning. *Educational Psychologist*, 40(4), 193–197. http://dx.doi.org/10.1207/s15326985ep4004_1.
- Azevedo, R., & Cromley, J. G. (2004). Does training on self-regulated learning facilitate students' learning with hypermedia? *Journal of Educational Psychology*, 96, 523–535. <http://dx.doi.org/10.1037/0022-0663.96.3.523>.
- (*)Ben-David, A., & Zohar, A. (2009). Contribution of meta-strategic knowledge to scientific inquiry learning. *International Journal of Science Education*, 31, 1657–1682. <http://dx.doi.org/10.1080/09500690802162762>.
- Benjamini, Y., & Hochberg, Y. (1995). Controlling the false discovery rate: A practical and powerful approach to multiple testing. *Journal of the Royal Statistical Society: Series B (Methodological)*, 57, 289–300.
- (*)Berkeley, S., Mastropieri, M. A., & Scruggs, T. E. (2011). Reading comprehension strategy instruction and attribution retraining for secondary students with learning and other mild disabilities. *Journal of Learning Disabilities*, 44, 18–32. <http://dx.doi.org/10.1177/0022219410371677>.
- (*)Blank, L. M. (2000). A metacognitive learning cycle: A better warranty for student understanding? *Science Education*, 84, 486–506. [http://dx.doi.org/10.1002/1098-237X\(200007\)84:4<486::AID-SCE4>3.0.CO;2-U](http://dx.doi.org/10.1002/1098-237X(200007)84:4<486::AID-SCE4>3.0.CO;2-U).
- Boekaerts, M. (1997). Self-regulated learning: A new concept embraced by researchers, policy makers, educators, teachers, and students. *Learning and Instruction*, 7, 161–186. [http://dx.doi.org/10.1016/S0959-4752\(96\)00015-1](http://dx.doi.org/10.1016/S0959-4752(96)00015-1).
- Boekaerts, M., Pintrich, P. R., & Zeidner, M. (Eds.). (2000). *Handbook of self-regulation*. San Diego: Academic Press.
- Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2009). *Introduction to meta-analysis*. Chichester, UK: Wiley. <http://dx.doi.org/10.1002/9780470743386>.
- Bransford, J., Brown, A. L., & Cocking, R. R. National Research Council, (U.S.). (1999). *How people learn: Brain, mind, experience, and school*. Washington, D.C: National Academy Press.
- (*)Broer, N. A., Aarnoutse, C. A. J., Kieviet, F. K., & van Leeuwe, J. F. J. (2002). The effect of instructing the structural aspect of texts. *Educational Studies*, 28, 213–238. <http://dx.doi.org/10.1080/0305569022000003681>.
- Brown, A. L., & Palincsar, A. S. (1989). Guided, cooperative learning and individual knowledge acquisition. In L. B. Resnick (Ed.). *Knowing, learning and instruction: Essays in honor of Robert Glaser* (pp. 393–451). Hillsdale, NJ: Lawrence Erlbaum.
- (*)Brunstein, J. C., & Glaser, C. (2011). Testing a path-analytic mediation model of how self-regulated writing strategies improve fourth graders' composition skills: A randomized controlled trial. *Journal of Educational Psychology*, 103, 922–938. <http://dx.doi.org/10.1037/a0024622>.
- (*)Carretti, B., Caldarola, N., Tencati, C., & Cornoldi, C. (2014). Improving reading comprehension in reading and listening settings: The effect of two training programmes focusing on metacognition and working memory. *British Journal of Educational Psychology*, 84, 194–210. <http://dx.doi.org/10.1111/bjep.12022>.
- Chiu, C. W. T. (1998). Synthesizing metacognitive interventions: What training characteristics can improve reading performance? *Paper presented at the annual meeting of the American Educational Research Association, San Diego, CA*.
- Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
- Cromley, J. G., Snyder-Hogan, L. E., & Luciw-Dubas, U. A. (2010). Reading comprehension of scientific text: A domain-specific test of the direct and inferential

- mediation model of reading comprehension. *Journal of Educational Psychology*, 102, 687–700. <http://dx.doi.org/10.1037/a0019452>.
- De Boer, H., Donker, A. S., & van der Werf, M. P. C. (2014). Effects of the attributes of educational interventions on students' academic performance: A meta-analysis. *Review of Educational Research*, 84, 509–545. <http://dx.doi.org/10.3102/0034654314540006>.
- Dent, A. L., & Koenka, A. C. (2016). The relation between self-regulated learning and academic achievement across childhood and adolescence: A meta-analysis. *Educational Psychology Review*, 28, 425–474. <http://dx.doi.org/10.1007/s10648-015-9320-8>.
- (*)Desoete, A., Roeyers, H., & De Clercq, A. (2003). Can offline metacognition enhance mathematical problem solving? *Journal of Educational Psychology*, 95, 188–200. <http://dx.doi.org/10.1037/0022-0663.95.1.188>.
- Dietrichson, J., Bøg, M., Filges, T., & Klitt Jørgensen, A.-M. (2017). Academic interventions for elementary and middle school students with low socioeconomic status. *Review of Educational Research*, 87, 243–282. <http://dx.doi.org/10.3102/0034654316687036>.
- Dignath, C., & Büttner, G. (2008). Components of fostering self-regulated learning among students. A meta-analysis on intervention studies at primary and secondary school level. *Metacognition Learning*, 3, 231–264. <http://dx.doi.org/10.1007/s11409-008-9029-x>.
- Dignath, C., Büttner, G., & Langfeldt, H. (2008). How can primary school students learn self regulated learning strategies most effectively? A meta analysis on self regulation training programmes. *Educational Research Review*, 3, 101–129. <http://dx.doi.org/10.1016/j.edurev.2008.02.003>.
- Donker, A. S., de Boer, H., Kostons, D., Dignath-van Ewijk, C. C., & van der Werf, M. P. C. (2014). Effectiveness of learning strategy instruction on academic performance: A meta-analysis. *Educational Research Review*, 11, 1–26. <http://dx.doi.org/10.1016/j.edurev.2013.11.002>.
- (*)Dresel, M., & Haugwitz, M. (2008). A computer-based approach to fostering motivation and self-regulated learning. *The Journal of Experimental Education*, 77, 3–18. <http://dx.doi.org/10.3200/JEXE.77.1.3-20>.
- OECD (2014). *Education at a glance 2014: OECD indicators*. OECD Publishing <http://dx.doi.org/10.1787/eag-2014-en>.
- (*)Festas, I., Oliveira, A. L., Rebelo, J. A., Damião, M. H., Harris, K., & Graham, S. (2015). Professional development in self-regulated strategy development: Effects on the writing performance of eighth grade Portuguese students. *Contemporary Educational Psychology*, 40, 17–27. <http://dx.doi.org/10.1016/j.cedpsych.2014.05.004>.
- (*)Fidalgo, R., Torrance, M., & Garcia, J. (2008). The long-term effects of strategy-focussed writing instruction for grade six students. *Contemporary Educational Psychology*, 33, 672–693. <http://dx.doi.org/10.1016/j.cedpsych.2007.09.001>.
- Flavell, J. H. (1979). Metacognition and cognitive monitoring. A new area of cognitive-development inquiry. *American Psychologist*, 34, 906–911. <http://dx.doi.org/10.1037/0003-066X.34.10.906>.
- (*)Gayo, E., Deaño, M., Conde, Á., Ribeiro, I., Cadime, I., & Alfonso, S. (2014). Effect of an intervention program on the reading comprehension processes and strategies in 5th and 6th grade students. *Psicothema*, 26, 464–470.
- (*)Glaser, C., & Brunstein, J. C. (2007). Improving fourth-grade students' composition skills: Effects of strategy instruction and self-regulation procedures. *Journal of Educational Psychology*, 99, 297–310. <http://dx.doi.org/10.1037/0022-0663.99.2.297>.
- (*)Hacker, D. J., Dole, J. A., Ferguson, M., Adamson, S., Roundy, L., & Scarpulla, L. (2015). The short-term and maintenance effects of self-regulated strategy development in writing for middle school students. *Reading & Writing Quarterly: Overcoming Learning Difficulties*, 31, 351–372. <http://dx.doi.org/10.1080/10573569.2013.869775>.
- Haller, E., Child, D. A., & Wallberg, H. J. (1988). Can comprehension be taught? A quantitative synthesis of “metacognitive” studies. *Educational Researcher*, 17, 5–8.
- Harackiewicz, J. M., Barron, K. E., Pintrich, P. P., Elliot, A. J., & Thrash, T. M. (2002). Revision of achievement goal theory: Necessary and illuminating. *Journal of Educational Psychology*, 94, 638–645. <http://dx.doi.org/10.1037/0022-0663.94.3.638>.
- Hattie, J., Biggs, J., & Purdie, N. (1996). Effects of learning skills interventions on student learning: A meta-analysis. *Review of Educational Research*, 66, 99–136. <http://dx.doi.org/10.3102/00346543066002099>.
- Hedges, L. V., & Olkin, I. (1985). *Statistical methods for meta-analysis*. San Diego, CA: Academic Press.
- (*)Heemsoth, T., & Heinze, A. (2016). Secondary school students learning from reflections on the rationale behind self-made errors: A field experiment. *The Journal of Experimental Education*, 84, 98–118. <http://dx.doi.org/10.1080/00220973.2014.963215>.
- (*)Hu, W., Wu, B., Jia, X., Yi, X., Duan, C., Meyer, W., et al. (2013). Increasing students' scientific creativity: The “learn to think” intervention program. *Journal of Creative Behavior*, 47, 3–21. <http://dx.doi.org/10.1002/jobc.20>.
- (*)Iseman, J. S., & Naglieri, J. A. (2011). A cognitive strategy instruction to improve math calculation for children with ADHD and ID: A randomized controlled study. *Journal of Learning Disabilities*, 44, 184–195. <http://dx.doi.org/10.1177/0022219410391190>.
- (*)Jitendra, A. K., Harwell, M. R., Dupuis, D. N., Karl, S. R., Lein, A. E., Simonson, G., et al. (2015). Effects of a research-based intervention to improve seventh-grade students' proportional problem solving: A cluster randomized trial. *Journal of Educational Psychology*, 107, 1019–1034. <http://dx.doi.org/10.1037/edu0000039>.
- (*)Jitendra, A. K., Hoppes, M. K., & Xin, Y. P. (2000). Enhancing main idea comprehension for students with learning problems: The role of a summarization strategy and self-monitoring instruction. *The Journal of Special Education*, 34, 127–139. <http://dx.doi.org/10.1177/002246690003400302>.
- Kostons, D., Van Gog, T., & Paas, F. (2010). Self-assessment and task selection in learner-controlled instruction: Differences between effective and ineffective learners. *Computers & Education*, 54, 932–940. <http://dx.doi.org/10.1016/j.compedu.2009.09.025>.
- Lan, Y.-C., Lo, Y.-L., & Hsu, Y.-S. (2014). The effects of meta-cognitive instruction on students' reading comprehension in computerized reading contexts: A quantitative meta-analysis. *Educational Technology & Society*, 17(4), 186–202.
- Lenhard, W., Baier, H., Endlich, D., Schneider, W., & Hoffmann, J. (2013). Rethinking strategy instruction: Direct reading strategy instruction versus computer-based guided practice. *Journal of Research in Reading*, 36(2), 223–240. <http://dx.doi.org/10.1111/j.1467-9817.2011.01505.x>.
- Lipsey, M. W. (2009). The primary factors that characterize effective interventions with juvenile offenders: A meta-analytic overview. *Victims and Offenders*, 4, 124–147. <http://dx.doi.org/10.1080/15564880802612573>.
- Lipsey, M. W., & Wilson, D. B. (2001). *Practical meta-analysis*. Thousand Oaks, CA: Sage Publications.
- (*)Mason, L. H. (2004). Explicit self-regulated strategy development versus reciprocal questioning: Effects on expository reading comprehension among struggling readers. *Journal of Educational Psychology*, 96, 238–296. <http://dx.doi.org/10.1037/0022-0663.96.2.283>.
- (*)Merchie, E., & Van Keer, H. (2016). Mind mapping as a meta-learning strategy: stimulating pre-adolescents' text-learning strategies and performance? *Contemporary Educational Psychology*, 46, 128–147. <http://dx.doi.org/10.1016/j.cedpsych.2016.05.005>.
- Mayer, R. E. (2008). *Learning and instruction* (2nd ed.). Upper Saddle River, NJ: Pearson Merrill Prentice Hall.
- (*)Ng, C. C., Bartlett, B., Chester, I., & Kersland, S. (2013). Improving reading performance for economically disadvantaged students: Combining strategy instruction and motivational support. *Reading Psychology*, 34, 257–300. <http://dx.doi.org/10.1080/02702711.2011.632071>.
- Palincsar, A. S., & Brown, A. L. (1984). Reciprocal teaching of comprehension-fostering and comprehension-monitoring activities. *Cognition and Instruction*, 1, 117–175. http://dx.doi.org/10.1207/s1532690xci10102_1.
- Peters, J. L., Sutton, A. J., Jones, D. R., Abrams, K. R., & Rushton, L. (2007). Performance of the trim and fill method in the presence of publication bias and between-study heterogeneity. *Statistics in Medicine*, 26, 4544–4562. <http://dx.doi.org/10.1002/sim.2889>.
- Pintrich, P. R. (2000). The role of goal orientation in self-regulated learning. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 451–502). San Diego, CA: Academic Press. <http://dx.doi.org/10.1016/B978-012109890-2/50043-3>.
- Pintrich, P. R. (2002). The role of metacognitive knowledge in learning, teaching, and assessing. *Theory Into Practice*, 41, 219–225. http://dx.doi.org/10.1207/s15430421tip4104_3.
- Pintrich, P. R. (2003). A motivational science perspective on the role of student motivation in learning and teaching contexts. *Journal of Educational Psychology*, 95, 667–686. <http://dx.doi.org/10.1037/0022-0663.95.4.667>.
- Pintrich, P. R. (2004). A conceptual framework for assessing motivation and self-regulated learning in college students. *Educational Psychology Review*, 16, 385–407. <http://dx.doi.org/10.1007/s10648-004-0006-x>.
- Pintrich, P., Smith, D., Garcia, T., & McKeachie, W. (1991). *A manual for the use of the motivated strategies for learning questionnaire (MSLQ)*. Ann Arbor, Michigan.
- Polanin, J. R. (2013). *Addressing the issue of meta-analysis multiplicity in education and psychology* Dissertations. Paper 539 http://ecommons.luc.edu/luc_diss/539.
- Pressley, M. (2002). Comprehension strategies instruction: A turn-of-the-century status report. In C. C. Block, & M. Pressley (Eds.), *Comprehension instruction: Research-*

- based best practices (pp. 11–27). New York, NY: Guilford.
- Raudenbush, S. W., Bryk, A. S., & Congdon, R. (2004). *HLM 6 for windows [computer software]*. Skokie, IL: Scientific Software International, Inc.
- (*)Sanz de Acedo Lizarraga, M. L., Sanz de Acedo Baquedano, M. T., Goicoa Mangado, T., & Cardelle-Elawar, M. (2009). Enhancement of thinking skills: Effects of two intervention methods. *Thinking Skills and Creativity*, 4, 30–43. <http://dx.doi.org/10.1016/j.tsc.2008.12.001>.
- (*)Sanz de Acedo Lizarraga, M. L., Ugarte, M. D., Iriarte, M. D., & Sanz de Acedo Baquedano, M. T. (2003). Immediate and long-term effects of a cognitive intervention on intelligence, self-regulation, and academic achievement. *European Journal of Psychology of Education*, 18, 59–74. <http://dx.doi.org/10.1007/BF03173604>.
- Schneider, W., & Pressley, M. (1997). *Memory development between two and twenty* (2nd ed.). Mahwah, NJ, US: Lawrence Erlbaum Associates Publishers.
- (*)Schünemann, N., Spörer, N., & Brunstein, J. C. (2013). Integrating self-regulation in whole-class reciprocal teaching: A moderator-mediator analysis of incremental effects on fifth graders' reading comprehension. *Contemporary Educational Psychology*, 38, 289–305. <http://dx.doi.org/10.1016/j.cedpsych.2013.06.002>.
- (*)Shayer, M., & Adhami, M. (2010). Realizing the cognitive potential of children 5-7 with a mathematics focus: Post-test and long-term effects of a 2-year intervention. *British Journal of Educational Psychology*, 80, 363–379. <http://dx.doi.org/10.1348/000709909X482363>.
- (*)Souvignier, E., & Mokhesgerami, J. (2006). Using self-regulation as a framework for implementing strategy instruction to foster reading comprehension. *Learning and Instruction*, 16, 57–71. <http://dx.doi.org/10.1016/j.learninstruc.2005.12.006>.
- (*)Spörer, N., & Schünemann, N. (2014). Improvements of self-regulation procedures for fifth graders' reading competence: Analyzing effects on reading comprehension, reading strategy performance, and motivation for reading. *Learning and Instruction*, 33, 147–157. <http://dx.doi.org/10.1016/j.learninstruc.2014.05.002>.
- (*)Stoeger, H., Sontag, C., & Ziegler, A. (2014). Impact of a teacher-led intervention on preference for self-regulated learning, finding main ideas in expository texts, and reading comprehension. *Journal of Educational Psychology*, 106, 799–814. <http://dx.doi.org/10.1037/a0036035>.
- (*)Teong, S. K. (2003). The effect of metacognitive training on mathematical word-problem solving. *Journal of Computer Assisted Learning*, 19, 46–55. <http://dx.doi.org/10.1046/j.0266-4909.2003.00005.x>.
- (*)Van Keer, H., & Verhaeghe, J. P. (2005). Effects of explicit reading strategies instruction and peer tutoring on second and fifth graders' reading comprehension and self-efficacy perceptions. *The Journal of Experimental Education*, 73, 291–329. <http://dx.doi.org/10.3200/JEXE.73.4.291-329>.
- Veenman, V. J., Van Hout-Wolters, B. H. A. M., & Afflerbach, P. (2006). Metacognition and learning: Conceptual and methodological issues. *Metacognition and Learning*, 1, 3–14. <http://dx.doi.org/10.1007/s11409-006-6893-0>.
- Weinstein, C. E., & Mayer, R. E. (1986). The teaching of learning strategies. In M. Wittrock (Ed.), *Handbook of research on teaching* (pp. 315–327). New York, NY: Macmillan.
- Wigfield, A., & Eccles, J. S. (2002). Expectancy value theory of achievement motivation. *Contemporary Educational Psychology*, 25, 68–81.
- (*)Wright, J., & Jacobs, B. (2003). Teaching phonological awareness and metacognitive strategies to children with reading difficulties: A comparison of two instructional methods. *Educational Psychology*, 23, 17–47. <http://dx.doi.org/10.1080/01443410303217>.
- Xin, Y. P., & Jitendra, A. K. (1999). The effects of instruction in solving mathematical word problems for students with learning problems: A meta-analysis. *The Journal of Special Education*, 32, 40–78.
- (*)Zepeda, C. D., Richey, J. E., Ronevich, P., & Nokes-Malach, T. (2015). Direct instruction of metacognition benefits adolescent science learning, transfer, and motivation: An in vivo study. *Journal of Educational Psychology*, 107, 954–970. <http://dx.doi.org/10.1037/edu0000022>.
- Zimmerman, B. J. (1986). Becoming a self-regulated learner: Which are the key subprocesses? *Contemporary Educational Psychology*, 11, 307–313.
- Zimmerman, B. J. (1990). Self-regulated learning and academic achievement: An overview. *Educational Psychologist*, 25, 3–17. http://dx.doi.org/10.1207/s15326985ep2501_2.
- Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory Into Practice*, 41, 64–70. http://dx.doi.org/10.1207/s15430421tip4102_2.
- (*)Zohar, A., & David, A. B. (2008). Explicit teaching of meta-strategic knowledge in authentic classroom situations. *Metacognition and Learning*, 3, 59–82. <http://dx.doi.org/10.1007/s11409-007-9019-4>.