Instructional Science (2019) 47:39–68
https://doi.org/10.1007/s11251-018-9475-z

ORIGINAL RESEARCH

The value of fixed versus faded self-regulatory scaffolds on fourth graders' mathematical problem solving

Stella Gidalevich[1,2,3] · **Bracha Kramarski**[3]

Received: 15 February 2018 / Accepted: 31 October 2018 / Published online: 30 November 2018

Abstract

Research has indicated that students can be taught self-regulated learning (SRL) in scaffolding programs focusing on a fixed continuous practice (e.g., metacognitive question prompts). However, the fading role of scaffolding to prepare autonomous learning is often an overlooked component. A unique approach for fading is suggested that offers a graduated reduction model of scaffolding prompts according to the SRL phases involved in the solution, which allows assimilation of processes to prepare learners for autonomous activity. This quasi-experimental study of fourth-graders ($n = 134$) examines the effectiveness of metacognitive self-question prompts in a *Fixed (continuous)* versus *Faded (graduated reduction*) scaffolds model during planning, monitoring and reflection phases, on the facilitation of students' SRL (metacognition, calibration of confidence judgment, motivation), and sense making of mathematical problem solving at the end of the program (short-term effect) and 3 months later (long-term/lasting effect). Findings indicated that the *Faded Group* performed best in the metacognition knowledge aspect, motivation in the performance goal approach increased and, in the avoidance, goal decreased. No differences were found between the groups on the regulation aspect and calibration of confidence judgment in the solution success. Additionally, the *Faded* Group outperformed the *Fixed* Group on sense making of problem solving. These findings were manifested particularly in the long-term effect. The study supports theoretical claims relating the role of fading scaffolds to increase students' autonomous SRL (metacognition, motivation) and improvements in sense making, particularly on the long-term retention effect.

Keywords SRL scaffolds · *Fixed* versus *Faded* prompts · Mathematical sense making · Fourth grade

✉ Stella Gidalevich
Stellagidl@gmail.com

Bracha Kramarski
Bracha.Kramarski@biu.ac.il

[1] Oranim Academic College, Kiryat Tiv'on, Israel

[2] Shaanan Academic College, Haifa, Israel

[3] Bar- Ilan University, Ramat Gan, Israel

Introduction

Self-regulation in learning (SRL) is considered critical for 21st century success, both academically and later in life (Pintrich 2000; Zimmerman 2008). Overall, self-regulation processes are built on both metacognitive and motivational strategies in a learning context that helps students to consider not only *what* and *how* they learn, but also *whether* their gains attain their goals (Moos and Ringdal 2012; Zimmerman 2008).

Fortunately, research has indicated that despite either complexity, SRL skills can be taught while pinpointing the vital role of scaffolding programs in promoting self-regulatory processes among students (Mevarech and Kramarski 2014; Tzohar-Rozen and Kramarski 2014). However, it is important to pay attention to *how* and *when* scaffolding should be *faded* or removed, when learners have acquired the desired skills or concepts after a fixed period of scaffolding and are expected to become autonomous learners who can implement these skills without scaffolds (Pea 2004; Puntambekar and Hubscher 2005). However, the fading conditions under which these skills are effective after removing the scaffolds have not been sufficiently investigated. Particularly lacking is a comparison between immediate effect at the end of the program (short term) and delayed effect (long-term) as a lasting effect (Ge et al. 2012).

This study addressed these issues by examining whether fixed (i.e., continuous) metacognitive scaffolds, compared to unique fading scaffolds related to graduated reduction, facilitate SRL and sense making of mathematical problem solving among young students (age 9–10). We now elaborate on the theoretical background underpinning the research variables assessed at the end of the program and three-months later.

SRL: self-regulated learning

SRL involves proactive, constructive processes that identify the three main components of SRL, cognition, metacognition and motivation (Pintrich 2000; Zimmerman 2000, 2008). Researchers indicate the efficiency of SRL in the context in which the learning takes place (e.g., Pintrich 2000; Zimmerman 2000). For the purpose of the current study, the content of sense making of mathematical problem solving was selected.

Cognition and metacognition components

Cognition is the process involved in knowing, that refers to the use of simple strategies like memorization, information processing, and higher-level strategies such as problem-solving and critical thinking. Metacognition is a form of cognition, a second or higher-order thinking process for understanding the task and the solution strategy (Flavell 1979; Pintrich 2000; Schraw 1998; Zimmerman 2008). Metacognition contains three aspects: knowledge of cognition, judgments and regulation of cognition (Pintrich et al. 2000). *Knowledge of cognition* is comprised of *declarative* knowledge about strategy/task ("what"?), *procedural* knowledge used on various cognitive strategies ("how"?), and *conditional* knowledge ("when"? and "why"?) that is important for the flexible and adaptive use of various cognitive strategies (Pintrich 2002; Schraw 1998).

Compared to the metacognitive knowledge aspect that has a static nature, metacognitive judgments are more process-related and reflect metacognitive awareness/experiences in ongoing metacognitive activities that individuals may engage in as they monitor and perform a task (Pintrich et al. 2000). Retrospective confidence judgements (CJ), refer to learners' confidence in their responses, which are mostly realized at the end of the learning process ("I'm sure I was right"; Mihalca et al. 2017; Roebers et al. 2014). Studies indicated a high tendency of overestimation in CJ among students (Labuhn et al. 2010). The current study focuses on bias of confidence judgments, that is, the difference between actual performance and the judged performance (henceforth calibration of CJ, Schraw 2009, see elaboration in the Method section). Previous research indicated that realistic CJs can be developed among younger children, aged between 9 and 10 years (e.g., Roebers et al. 2014).

Similar to metacognitive judgments, the *regulation of cognition* refers to a dynamic strategic processing activity (Pintrich et al. 2000) that helps control one's thinking or learning (Nelson and Narens 1994) through three phases. In the planning (forethought) phase, learners set goals for their own planning of specific activities, to complete a specific task. Next, in the monitoring (performance) phase, learners use their goals to monitor the process and move it along to use goals as checkpoints for progress along tasks. Finally, in the reflection (evaluation) phase, learners use information gained from the completed task to improve performance on the next task, what worked, what did not work and *why* (Pintrich 2000; Zimmerman 2008). The knowledge of cognition belongs predominantly to domain-specific strategies, and regulation of cognition is domain-general for processing strategic activation and application via the three phases (Pintrich 2000; Schraw 1998; Veenman et al. 2006). Research literature indicates that SRL is not achieved naturally, so students should be motivated to engage in a particular task and to be strategic in that process that can influence different processes of self-regulation (Dignath et al. 2008; Meece 1994).

Motivation component

A general assumption shared by most models of self-regulated learning is that there is some type of goal, criterion, or standard against which to assess the operation of the learning process and which guides regulatory processes (Boekaerts and Corno 2005; Winne 1996; Zimmerman 2000). It is assumed that learners set standards or goals to strive for in their learning, monitor their progress toward these goals, and then adapt and regulate their cognition, motivation and behavior in order to reach these goals (Pintrich 2000). The Achievement Goals Orientation Theory which has emerged in recent years as one of the dominant theories of academic motivation linked directly to the SRL theory (Efklides 2011; Pintrich 2000; Wolters 2004; Zusho and Edwards 2011). The theory focuses on the reasons/goals students have for engaging or achieving in a learning situation (Ames 1992; Dweck 1986; Elliot 1997; Kaplan and Maehr 2002). This theory originally stressed two general orientations to achievement: mastery and performance goals (Ames, 1992; Dweck and Leggett 1988).

The present study adopted the basic distinction as suggested by Ames (1992) and Dweck (1986), namely, mastery and performance goal-orientations that function in an achievement situation. The primary difference between these two types of goal-orientations is whether learning is valued as an end in itself, or as a means to reach some external goals. In particular, students with mastery goals value learning for its own sake and prefer situations where they can expand new skills and gain new knowledge (Ames 1992). Mastery goals assist learners to improve their performance, advance their learning, and achieve deep level strategy use (Nolen 1988). On the other hand, students with performance-approach goals value ability and prefer situations where they can demonstrate their ability and compare it with other students (Ames 1992). Students with performance-avoidance goals aim to avoid exhibiting lack of ability, invest minimal effort and use surface strategies.

Many studies found that mastery-oriented students report higher levels of effective cognitive and metacognitive strategies (Pintrich 2000). Planning, monitoring and regulatory strategies, were associated with mastery goal-orientation (Al-Harthy and Was 2010; Chatzistamatiou et al. 2015; Vrugt and Oort 2008; Wolters 2004). However, the link between performance goals and SRL is less straightforward. Most previous studies examining the relation between goal-orientation and learning strategies focused on adolescent and college students (e.g. Weinstein et al. 2011; Wolters 2004). There is some evidence regarding the relations between goal-orientations and self-regulation strategies in elementary schools (e.g., Patrick et al. 2007), indicating that mastery goals were strongly related to self-regulation strategies used by fifth-grade students. Additionally, Seo and Kim (2001) found that in elementary school students, both mastery and performance-approach goals related positively to metacognitive strategies as well as to mathematics achievement. However, further research with a focus on younger students with SRL scaffolds (fixed vs. faded) is needed in order to understand better the above relations at earlier stages of SRL development in the context of mathematical problem solving.

Sense making of mathematical problem solving: short and long-term effects

Mathematical problem solving is known to be one of the most difficult topics for students to grasp (e.g., Mevarech and Kramarski 2014; Schoenfeld 1992; TIMSS 2011). This ability is defined as a *sense making* process that demands "understanding of a situation, context, or concept by connecting it with existing knowledge or previous experience" (NCTM 2009, p. 4). Studies have indicated that despite numerous efforts over the years to train students in sense making of mathematical problem solving (henceforth "sense making"), many learners of different ages still have difficulty in internalizing these processes for immediate use (short term effect), and in transferring them over time after long-term practice, thus indicating a lasting effect (Schoenfeld 1992; TIMSS 2011; Watts et al. 2016). These difficulties are explained by numerous researchers as resulting from students' limited cognitive information-processing, due to a lack of brain working memory (WM) space for new knowledge acquisition, that increases students' cognitive load and affects learners' mental effort (Ariës et al. 2014; Baddeley 1992; Goldberg 2010; Sweller 1988).

Research findings show that decreasing cognitive load can be attained by increasing learners' ability to self-regulate their knowledge and effort during matching between SRL and solution phases (Ariës et al. 2015; Goldberg 2010; Kramarski and Fridman 2014). This raises the need to investigate the desired SRL scaffolds effects and its fading conditions, in the short (immediate) and long terms (i.e., lasting effect).

SRL fixed versus faded scaffolds

The research literature contains diversified SRL scaffolding programs recommending explicit scaffolds and self-regulation strategy modeling in the mathematical context, which encourage autonomous learning, and then lead to improved SRL metacognition and goal orientation motivation (e.g., Adler et al. 2016) and mathematical achievements (e.g., Mevarech and Kramarski 2014).

Fixed scaffolds are a continuous "forced" method that allows practicing metacognitive scaffolds directed at the proper moment to each SRL phase, while solving various problems. The research generally supports effects of continuous scaffolds "as a catalyst" for eliciting self-regulation strategies and enhancing transfer performance in the context of mathematical problem solving (e.g., Kramarski et al. 2013; Hoffman and Spatariu 2008). Furthermore, whereas some studies found that metacognitive scaffolds reduce students WM overload in problem solving (e.g., Ariës et al. 2015; Kramarski and Fridman 2014), other researchers argue that a continuous repetition option in metacognitive scaffolds slows their autonomy in solving the problems alone without help, which may add to students' cognitive WM overload (Kramarski and Fridman 2014; Kramarski et al. 2013; Bannert and Mengelkamp 2013; Mayer 2009). Therefore, studies suggest that an effective metacognitive scaffold program should offer students some fading (reduction) of the scaffolds (Pea 2004; Puntambekar and Hubscher 2005).

Fading scaffolds is an extremely important method aspect of scaffolding, relating to how the supports are faded over time (Pea 2004; Puntambekar and Hubscher 2005), but is often a component missing in intervention programs. Most fading approaches relate to fading support all at once at the end of the program, while an advanced model for fading offers graduated reduction of skills to enhance students' autonomous activity in attaining the desired skills/aspects (e.g., Puntambekar and Hubscher 2005). The initial practice of the fading graduated reduction model contains detailed support (full aspects/skills) and then involves lessened support over time (Bulu and Pedersen 2010; McNeill et al. 2006). However, there are still open questions about what, when and how-to fade out (e.g., Jaakkola and Veermans 2018; Tawfik et al. 2018).

Studies on fixed versus faded scaffolds effects

Effects of the fixed and faded scaffolds have led to contradictory findings. For instance, Lee and Songer (2004) by intervention support for domain-specific scaffolding in technology enhanced science, compared fixed and graduated faded scaffolds (i.e., withdrawn along the curriculum) provided by three kinds of explanation (exemplars, questions, and sentence), among primary school students (fifth and sixth grade). They found that while learners in both groups exhibited a gain in content knowledge, the fixed group outperformed the faded group in writing scientific explanations from authentic data, which is a higher order thinking skill.

Bulu and Pedersen (2010) compared fixed and graduated faded scaffolds in the context of hypermedia science learning (sixth grade). Three types of scaffolds, examples, question prompts and sentence starters, were presented to the students in both groups (fixed vs. faded). Over the course of the program, scaffolds were withdrawn gradually from the faded group, one at a time, starting from the withdrawal of examples, then question prompts, and finally sentence starters. The study found that students with fixed scaffolds performed better than those with faded scaffolds, in terms of developing solutions and making justifications.

Jaakkola and Veermans (2018) investigated the effects of concreteness fading on learning and transfer across three grade levels (4–6) in elementary school science education, in comparison to learning with fixed (constant) concrete representations with a computer-based simulation environment. The study found a significant interaction between condition and grade level in relation to learning outcomes, suggesting that the outcomes generally improved as a function of grade level. It was found that learning with fixed concrete representations either took less time or resulted in better learning compared to concreteness fading. A surprising finding was that the concrete condition succeeded at least as well as the fading condition on transfer tasks.

Tawfik et al. (2018) compared the effect of *fixed* (sustained) versus *graduated faded* scaffolds with four kinds of prompts: Problem-representation, developing solutions, making justifications, and monitoring evaluation at the end of a computer environment program on undergraduate students' argumentation in ill-structured problem-solving. They found that students in the fixed (sustained) scaffolding condition performed significantly better than the students in the faded condition in scientific argumentation.

In contrast, McNeill et al. (2006) when comparing fixed versus graduated faded scaffolds in supporting students' construction of scientific explanations, found that students (seventh grade) who underwent graduated faded scaffolding gained better comprehension in scientific explanation (i.e., claim, evidence, and reasoning) in the post-test than students who were exposed to the fixed scaffolds. They concluded that fading written scaffolds better equipped students to write explanations when they were not provided with support.

Similarly, Kester and Kirschner (2009) investigated fixed versus graduated faded scaffolds in two types of support, conceptual (concept map) and strategic (flow chart) in a student-centered e-learning environment to help adult learners to reach hypertext navigation accuracy skills. It was hypothesized that fading support during practice helps learners navigate more accurately during practice and achieve a higher practice and test performance as compared to learners receiving full support or no support during practice. The study confirmed the beneficial effects of fading support on navigation accuracy, but not on effects of fading in similar practice tasks during the program and transfer items (unfamiliar test items) in a regulate written format (no navigation).

To summarize, findings regarding effects of fixed versus faded scaffolds are inconclusive. Studies among primary school students (fourth-seventh grade) have been conducted mainly in the science domain, focusing on the cognitive level of the strategy scaffold but not on the mathematics domain. These studies tested metacognitive scaffolds (fixed and faded) and their effect on both SRL components: metacognition and motivation (Pintrich 2000; Zimmerman 2000). To gain better insight into the effects of these two scaffold types s, this study investigates the development of SRL and problem solving as a consequence of a metacognitive prompt set serving as an instructional scaffold to foster SRL.

Prompts as an instructional scaffold for fostering SRL

Prompts are external stimuli questions or statements that indicate to learners when and how to engage in a productive processing engagement, with the objective of activating metacognition and motivational regulation processes (Bannert and Reimann 2012; Davis 2003; Panadero et al. 2012; Müller and Seufert 2018). Prompts are powerful instructional means to activate learners' existing self-regulation skills, by providing the balance between necessary external support and desired internal regulation (Mevarech and Kramarski 1997, 2014; Koedinger and Aleven 2007). Based on SRL theoretical relations and the effects found, prompts should not only increase the specific self-regulatory processes that are prompted as metacognition, but should also affect learners' perceptions of complementary components as motivation (Panadero et al. 2012; Pintrich 2000). Metacognitive prompts may lead learners to reflect on their skills and resources. As a consequence, learners might feel equipped and motivated for learning in an environment that requires learners' self-regulation when problem solving. Additionally, prompts often lead to learning success, which reflects mastery experiences (Devoldr et al. 2012; Zheng 2016), providing a major source for perceptions of motivational achievement goals (Ames 1992).

Scaffolding prompts can differ in their delivery mode and may be embedded or non-embedded, fixed or adaptive, direct or indirect (Devoldr et al. 2012; Zheng 2016). Embedded scaffolds are integrated in the learning environment so that students are obliged to pay attention to them, while non-embedded scaffolds are initiated by students themselves (Narciss et al. 2007).

The current study addresses the issue of fostering SRL (metacognition and motivation) with IMPROVE metacognitive self-question prompts oriented to *what*, *when*, *how* and *why* to use or withdraw scaffolds, and the effect of fading versus fixed scaffolds (Pea, 2004; Puntambekar and Hubscher 2005) in a unique faded reduction model, by focusing on younger students (grade four).

IMPROVE metacognitive self-question prompts in fixed versus faded scaffolds

The IMPROVE self-questioning method for mathematical teaching is one of the known methods for cultivating the learner's SRL phases to enhance mathematical problem solving. This model (Kramarski and Mevarech 2003; Mevarech and Kramarski 1997, 2014) aims at scaffolding key aspects of sense making in problem solving by using generic, self-directed question prompts: Comprehension, Strategy and Reflection oriented to the three SRL phases of planning, monitoring, and reflection (see Fig. 1).

IMPROVE metacognitive prompts explicitly directed to the SRL phases have the potential to promote any of the metacognitive processes (planning, monitoring, reflection), to affect goal-orientation motivation that reflects "… ways of approaching, engaging and representing to achievement situations" (Ames 1992, p. 261). Previous studies conducted mainly in secondary schools with IMPROVE prompts on a fixed continuous scaffolding period in mathematics, were successful in metacognition and in sense making of problem

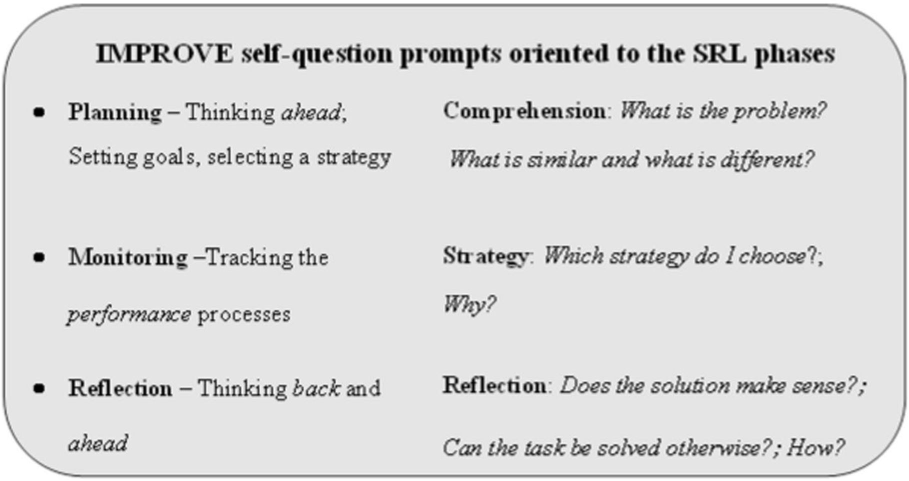


Fig. 1 The IMPROVE metacognitive prompts oriented to the SRL phases: planning, monitoring and reflection

solving (Kramarski and Mevarech 2003; Kramarski et al. 2013; Mevarech and Kramarski 1997, 2014), and showed a decrease in students' cognitive problem solving load (Kramarski and Fridman 2014).

However, these powerful effects were compared to control groups with no scaffolds (Mevarech and Kramarski 2014). Nevertheless, little research has been done to understand the mechanism relating to *when* and *how* to withdraw metacognitive scaffolds to encourage student's autonomy to set motivational *individual mastery goals,* and to enhance sense making of mathematical problem solutions after fading the prompting support, particularly for long-term effects.

Study design and the goals

This study examines the relative effectiveness of IMPROVE metacognitive self-questioning prompts in facilitating young students' SRL (metacognition, motivation), under the fixed versus faded (graduated reduction) scaffolding conditions and relating to sense making in a problem-solving context. The Fixed Group was exposed explicitly to the IMPROVE metacognitive prompts (comprehension, strategy, reflection) oriented to SRL phases (planning, monitoring, reflection) in a continuous scaffold (Fig. 1). The Faded Group was also exposed to the same prompts, phases and problems, but used them in accordance with the unique faded graduated reduction model (see elaboration in the Method section and Fig. 2).

The study examines the following questions:

Does the use of IMPROVE metacognitive self-question prompts with different levels of scaffolds: Fixed versus Faded, with the graduated reduction model, have a short-term effect on the students at the end of the program, and after a fading period of the metacognitive scaffolds (three-months, long-term lasting effect) with regard to-

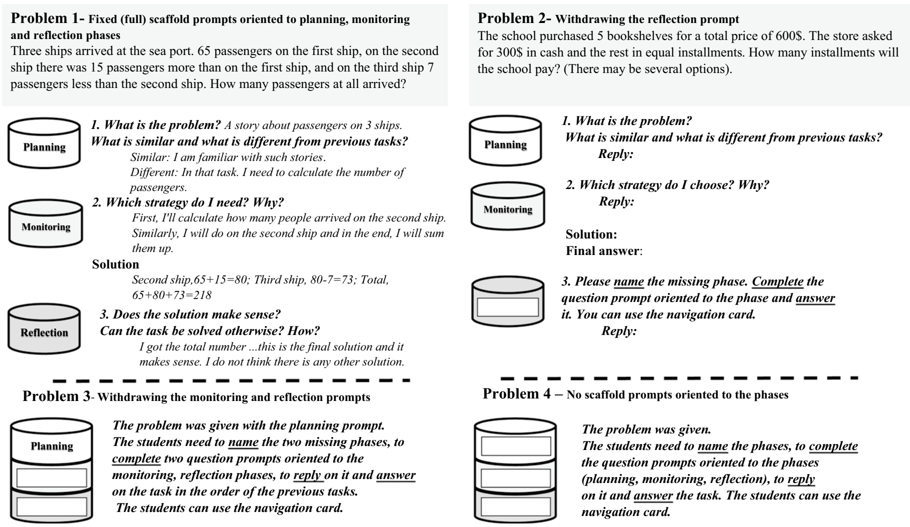


Fig. 2 Fixed versus Faded question prompts oriented to SRL phases embedded in a set of four sense making tasks

a. **SRL**: metacognition, CJ and motivation?
b. **Sense making** of problem solving?

Several theoreticians note that the graduated fading model provides optimal autonomous support to the learner in various domain content aspects and skills (Ge et al. 2012; Kester and Kirschner 2009; McNeill et al. 2006; Pea 2004; Puntambekar and Hubscher 2005). In this manner, it is expected that the Faded Group will assimilate the SRL phases more than the Fixed Group, thereby contributing to internalization of the metacognitive process, and sense making of mathematics, particularly in the long-term effect (Mevarech and Kramarski 2014; Kester and Kirschner 2009; Lee and Songer 2004; Schonfeld 1992). This is in line with research showing a greater impact on metacognitive effects in follow-up learning sessions than measured earlier in short term effects (Bannert et al. 2015).

Taking into account the theoretical relatedness between the SRL components (metacognition and motivation) we expect to find effects of the IMPROVE prompts also on students motivation (achievement goal orientation) in the *Faded* group compared to the *Fixed* group. Following the mixed effects research findings about the two scaffolding programs (fixed and faded), direct scaffold prompts on metacognition and not on motivation, and the young age of the participants (grade four; 9–10), we didn't formulate an explicit hypothesis regarding the effects of the two scaffolding types on SRL self-reported measures across the three intervals (T1, T2, T3).

Method

Participants

The current quasi-experimental study included 134 students from 4 classes (grade 4, ages 9–10) who attended two schools selected randomly in the same city in northern Israel with mid-level socioeconomic status, that were similar in their achievements on mathematics standardized measures tested by the Ministry of Education. Two classes, both fixed or faded with all their students (~33 students per class) in each school were randomly assigned into one of the two study groups: The Fixed Group (68 students; 50% boys and 50% girls), and the Faded Group (66 students; 51,5% boys and 48.5% girls). Assigning the two classes to one study group was done in order to minimize the confounding variable of the study groups. The two study groups did not differ by gender, χ^2 (1) = .03, p = .86; pretests as presented in the results section and teachers' professional background are presented in the training section. In particular, the two groups are similar in their class heterogeneity (SD; Cohen's d) on each of the measured variables. The study was approved by the Chief Scientist for Operation at the school.

Measures

Measures on SRL aspects (metacognition, calibration of CJ, motivation) and sense making of mathematics were administrated three times: Pre-intervention (T1), post-intervention (T2) for short-term (immediate) effect and after 3 months (T3) to assess long-term lasting effect.

Metacognition

The self-report Junior Metacognitive Awareness Inventory developed by Sperling et al. (2002) was administrated. The questionnaire contained 24 items, which produced two principal factors (Brown 1987): *Knowledge* of cognition (10 items, e.g., *"I know when I understand something"*) and *regulation* of cognition (14 items, e.g., *"I find myself pausing regularly to check my understanding"*) scored on a five-point Likert scale ranging from 1, "Never" to 5, "Always". Cronbach's Alpha reliability of the 24 items was $\alpha = .89$, $\alpha = .76$ for the knowledge of cognition and $\alpha = .81$ for the regulation of cognition.

Calibration *of CJ*

A "self-estimation ruler" of calibration of metacognitive CJ was presented at the end of the students' mathematical problem solution (see Appendix B). The ruler was integrated into the test at three-time points (T1, T2 and T3), with two problem-solving tasks in each test to examine the effects of the two interventions on participants' CJ. Participants were asked to estimate their success in reaching the solution with the following instruction: "Indicate on the ruler the degree of your success in solving the problem". The ruler was developed by Huff and Nietfeld (2009) and used a scale ranging from 0 (low estimation) to 100 (high estimation). The participants were given an explanation of the self-estimation role as part of metacognition self-awareness (what and why), and a demonstration of how to mark the judgments on the ruler. They practiced their CJ on a lesson before the beginning of the intervention, after solving a task and then compared their estimation with the correct solution. The analysis was conducted using a calculation calibration bias according to the third formula of Schraw (2009).

$$\text{Bias Index} = \frac{1}{N}\sum_{i=1}^{N}(c_i - p_i)$$

Calibration bias is the difference between the learner's self-estimation (i.e.,c_i scores) in executing the task and the actual performance (i.e., p_i*scores*). Accordingly, a difference aspiring to zero indicates a high level of accuracy, while a difference distant from zero towards the positive or negative direction, indicates inaccuracy of metacognitive judgment ability. A positive difference indicates a higher CJ than the actual score (given by the teacher), whereas a negative difference indicates a low CJ, meaning the learner's self-estimation was lower than the real score (Schraw 2009). Cronbach's Alpha reliability relating to calculation of the calibration bias was $\alpha = .76$.

Motivation

A self-report questionnaire based on *achievement goals* theory was administered at three time points (T1, T2 and T3). The questionnaire was developed by Midgley and colleagues (Midgley et al. 2000), including 19 statements relating to achievement goals, and used a five-point Likert scale ranging from 1, "Not true at all" to 5, "Very true." The respondent ranked each statement according to how true it was for him or her when solving sense making problems. Statements were divided into three groups: six statements about mastery goals (for example, "*It is important for me to understand clearly what I learn in class*"), where an answer of 1 showed a low mastery-goals orientation and an answer of 5 indicated a high mastery-goals approach; six statements relating to performance-approach goals (for example, "*It is important for me that students in my class think that I am good at tasks*"), where 1 showed low performance-approach goals and 5 indicated high performance-approach goals; and seven statements relating to performance-avoidance goals (for example: "*It is important for me not to look stupid at school*"), where 1 indicated low performance-avoidance goals and 5 indicated high performance-avoidance goals, which is less effective for self-regulation processes. The measure of performance-avoidance goal is ***inverted***; thus, *higher* scores indicate *less avoidance*. Cronbach's Alpha reliability of the 19 items was $\alpha = .88$, $\alpha = .80$ for the *mastery goals*, $\alpha = .89$ for the *performance-approach goal,* and $\alpha = .77$ for the *performance-avoidance goals.*

Sense making of mathematics

Tasks were taken from the standardized Meitzav exam for the fourth grade, developed by the Israeli Ministry of Education (2005, Version A). At each time point (T1, T2 and T3) the test consisted of 7 tasks on the same level of complexity (routine and novel), adapted by changing the numbers in each time version (see Appendix B). The scores ranged from 0 to 100 where "100" was awarded for a correct solution with a correct solution path that indicates their thinking strategy; "50" for a partial solution without the solution path or a wrong solution path, and "0" for a wrong answer without a solution path or a wrong solution path. The tests and indicators for the task were approved by the supervisor of mathematics teaching at the Ministry of Education and were recommended for use with fourth graders. Cronbach's Alpha reliability for the sense making test was $\alpha = .75$.

Intervention program structure

Teachers' background and instruction

The four female teachers who taught the participants each held a bachelor's degree in mathematics education. They have also similar experience in professional development programs for mathematics teaching methods in primary schools to enhance students' meaningful learning. Each teacher had more than 10 years of math teaching experience and was considered by the principal to be an expert teacher in the school according to official tests and surveys of student satisfaction.

Teachers were instructed in each school by the first author on their intervention program structure (6 h'; Kramarski 2017; Kramarski and Revach 2009). Instructions was supported with written protocols, scripts for presenting the theoretical SRL framework, rationale and techniques for using and discussing it in class (see Fidelity section below for details). Teachers were informed that they were participating in an experiment (12 lessons in 7 weekly sessions) in which a new program and materials were being designed for teaching mathematics based on theories about SRL, autonomous learning in the context of sense making in mathematics.

For both groups, the first part of the instruction focused explicitly on the rationale for SRL (metacognition, CJ, motivation) and its phases (Pintrich 2000; Zimmerman 2008), using the IMPROVE metacognitive self-question prompts (comprehension, strategy and reflection; Mevarech and Kramarski 1997, 2014). In the second part, teachers were exposed explicitly to the rationale, techniques and modeling of their program (fixed or faded).

Teachers practiced the problem-solving process with a set of the 4 sense making tasks with the IMPROVE prompts along the SRL phases (fixed or faded) as presented below in Fig. 2 and discussed possible difficulties that students could face in implementing such a model related to their answers on the prompts, and on the sense making of the solution.

Finally, a short explanation was given regarding the testing measures (SRL, CJ, sense making), students' folders, students' navigation cards and how to summarize the program in the last meeting as presented on the time line of the program (Fig. 3).

Fixed versus faded groups structure

Each intervention lasted for 7 weekly sessions, including 12 lessons (the first-introduction and last-summary session included one lesson and 5 sessions included double lessons). The duration of the program's problem-solving structure is set in accordance with the scope of hours allocated to learning sense making of problem solving in the national curriculum for grade 4. The detailed structure of each session is presented on Fig. 2. Students in both groups in each session solved 4 mathematical sense making problems at the same level of complexity related to the official curriculum scaffold by the IMPROVE self-question prompts oriented to the SRL phases (planning, monitoring, reflection). The last session practice was based on higher complex sense making problems (see example type on Appendix B).

The *Fixed* Group was exposed in each problem to the full IMPROVE self-question prompts (comprehension, strategy, reflection; Fig. 1) oriented respectively to each SRL phase (planning, monitoring, reflection). Students were asked to answer in the prompts by writing and solving the problem (see Fig. 2; problem 1 for a full prompts solution process). The *Faded* Group was prompted according to the graduated fading model as follows (see elaboration in Fig. 2):

- The *first* problem was presented similarly to the *Fixed* Group, with the ***three*** metacognitive prompts ***fully*** oriented respectively to each of the three phases. Students were asked first to reply to the planning and monitoring question prompts by writing, to solve the problem and finally, to reply to the reflection prompt.

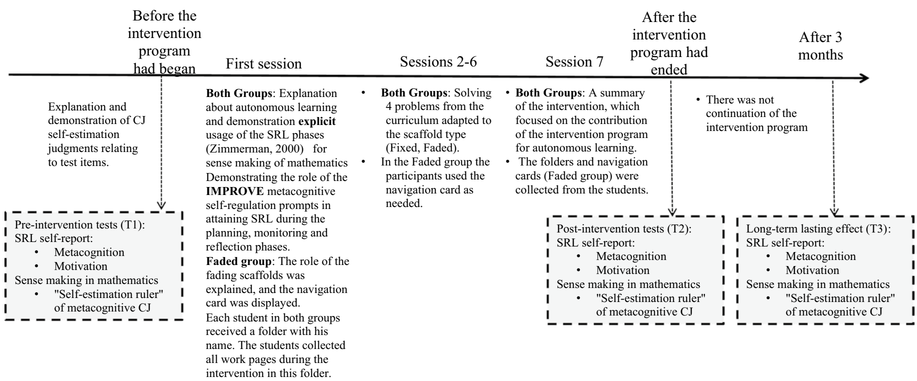


Fig. 3 Overview of the *Fixed versus Faded* groups' study program and timeline

- The *second* problem was prompted with *two* questions oriented respectively to each of the two phases: *planning* and *monitoring*; students were asked to reply to the two question prompts by writing and by solving the problem. Finally, they named and wrote out the withdrawn question prompt for *reflection* and replied to it.
- The *third* problem was prompted with *one* question oriented to the planning phase, and students were asked to name and write out the two withdrawn question prompts for *monitoring* and for *reflection.* Then, to reply to the question prompts by writing and to solve the problem in the same order as explained in the previous problems.
- The *fourth* problem was *not* prompted, but was solved autonomously by students, while naming and writing out the withdrawn question prompts oriented to the three phases, replying to the question prompts and solving the problem in the same order as they had done previously.

Students then took turns to read their answers aloud to the self-questions prompts and problem solution after solving it. The teacher led a class discussion about these answers and wrote examples on the board. Each student collected his work pages in a personal folder. The teacher kept the folders in the classroom during the intervention and presented them to the researcher in his program fidelity visit. In order to help the students in both groups to assimilate the metacognitive questions and phases, they were given a *navigation card* as presented in Fig. 1, with all of the self-questions to use during the program. The same duration time for solving the problem was given in both groups.

Fidelity of the program

During the course of the intervention, the first author observed each second lesson in both groups. In each observation, the researcher rated the extent (1-low to 4-high) to which the instructor implemented the intervention precisely according to the program's scripts, tasks, personal folders; procedures, instruction, metacognitive prompts, and discussion time allocation, according to the goals of the study group. Overall, implementation fidelity was high (for the total observations: $M = 3.93$, $SD = 0.29$, range: 3–4). Deviations were minor, infrequent, and easily corrected (e.g., only on two occasions was time allocation rated less than 4 for discussion time allocation—rated 3 once for one Fixed Group teacher and once for one Faded Group teacher). At the end of each lesson, the researcher gave the teacher feedback regarding the observed criteria.

Results

A preliminary MANOVA conducted on the pretest measures indicated no differences between the groups on SRL aspects (metacognition, CJ, motivation) and sense making of problem solving. Thus, mixed ANOVA with repeated measures analyses on time and on group were conducted on the post test. In each analysis, the *group* (Faded and Fixed) was treated as the between-subjects variable and the *time* (T1, T2, T3) as the within-subjects variable. Mauchly's sphericity test was conducted to test the assumption regarding the condition where the variances of the differences between all possible pairs of within-subject conditions (i.e., levels of the independent variable) are equal. The test result was significant, and we reported the adjustment results from the

Greenhouse–Geisser test. The design was balanced so that the number of observations in each cell was equal, thereby avoiding any violation of the homogeneity assumption.

Metacognition

To examine scores in the metacognition scales (Schraw and Dennison 1994): *knowledge* of cognition and *regulation* of cognition, two 2×3 repeated measures analyses of variance (mixed ANOVA) were conducted.

Knowledge of cognition

The two way *interaction* of group x time was significant, $F(2, 131) = 4.26$, $p < .05$, $\eta_p^2 = .06$ (Table 1). Simple effect analyses comparing the three time points for each study group separately indicated significant differences between the three time points in both groups, [$F(2, 64) = 4.23$, $p < .05$, $\eta_p^2 = .12$ for the *Faded* Group and $F(2, 66) = 5.12$, $p < .01$, $\eta_p^2 = .13$ for the *Fixed* Group]. Bonferroni analyses indicated that while *higher* scores were found in T2 and T3 compared to T1 ($d = 0.30$; $d = 0.34$ respectively) in the *Faded* Group ($p < .01$), with no significant differences in scores between T2 and T3 ($p = .86$), a significant *decrease* in scores was found in T3 compared to T2 ($d = -0.37$) in the *Fixed* Group ($p > .05$), but no significant differences in scores were found between T1 and T2 ($p = .13$) or between T1 and T3 ($p = .99$). Comparing the two study groups at each time point reveals that while no significant differences were found between the two groups in T1 ($p = .74$) and T2 ($p = .98$), higher scores in the *Faded* Group, compared to the *Fixed* Group, were found in T3 ($p < .05$, $d = 0.43$) (Fig. 4).

Table 1 Means (and SD) and *F*-score of the metacognition and sense making variables by groups: fixed ($n = 68$) and faded ($n = 66$) and three-time points

Dependent variables	Study groups	Pre (T1)		Post (T2)		Long-term (T3)		F-Score	
		M	SD	M	SD	M	SD	F Time	F Interaction
Knowledge of cognition[1]	Fixed	4.01	0.59	4.15	0.59	3.96	0.63	5.39**	4.26*
	Faded	3.97	0.78	4.15	0.69	4.23	0.64		
Regulation of cognition[1]	Fixed	3.82	0.60	3.93	0.65	3.81	0.66	2.12	1.29
	Faded	3.81	0.78	3.93	0.69	3.97	0.83		
Calibration of Confidence Judgements[2]	Fixed	14.52	38.81	42.18	36.76	4.49	29.03	50.03***	.43
	Faded	14.24	31.43	36.74	41.69	– 2.99	25.05		
Sense making	Fixed	66.91	28.07	61.50	26.57	72.00	23.01	22.22***	3.04*
	Faded	63.89	29.75	64.65	24.90	80.40	18.41		

[1]Five-point Likert scale ranging from “1”: Never, to “5”: Always

[2]Was conducted using a calculation calibration bias: the gap between the learner’s self-estimation in executing the task and the actual score (the third formula, Schraw 2009)

*p < .05, **p < .01, ***p < .001

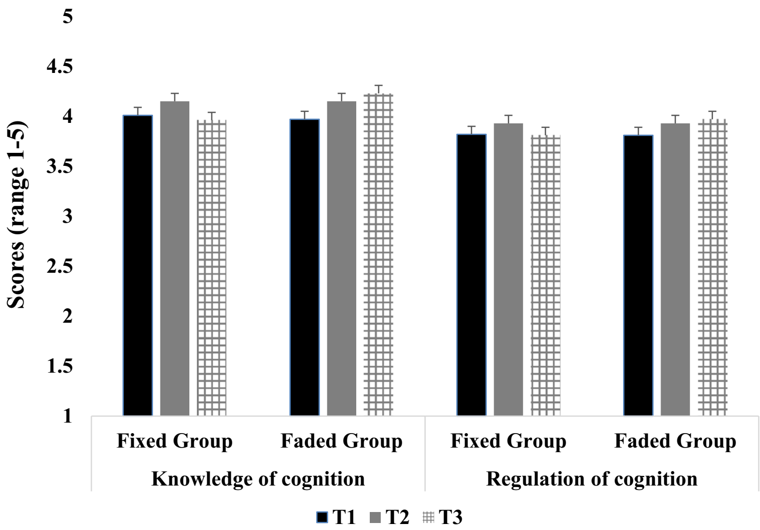


Fig. 4 Metacognition scores of two study groups for three-time points

Regulation of cognition

Data in Table 1 show slight differences between the three-time measures in each group. Both groups *increased* their self-report *regulation* of cognition ability at the end of the program (T2; $d = 0.18$). Whereas the *Fixed* Group *decreased* that measure on *lasting* effect (T3; $d = -0.18$), the *Faded* Group maintained the same achievement ($d = 0.06$). However, repeated measures analyses of variance (mixed ANOVA) indicated that the two way *interaction* of group x time was also not significant, $F(2, 131) = 1.29$, $p = .28$, $\eta_p^2 = .02$ (Table 1).

Confidence judgment

To examine scores in the calibration of confidence judgment, a 2x3 repeated measures analysis of variance (mixed ANOVA) was conducted. A main effect of *time* was found, $F(2, 131) = 50.03$, $p < .001$, $\eta_p^2 = .43$, indicating differences in T3 compared to T1 ($p < .01$) and T2 ($p < .001$) ($d = 0.30$ and $d = 0.87$, respectively), and in T2 compared to T1 and T3 ($ps < .001$) ($d = -0.47$, $d = -0.87$, respectively). In continuous to over confidence judgments' bias on T1, both groups (*Fixed & Faded*) further increased on T2 their over confidence judgments' bias of their solution success on the short—term effect (end of the program), whereas on the long-term effect in T3 (3-months later) their confidence judgment was more realistic. In T3, the Fixed Group had a higher index bias score (M = − 4.49) compared to the Faded Group (M = 2.99) indicated greater accuracy in self-estimation. Finally, the main effect of *group*, $F(1, 132) = 1.60$, $p = .21$, $\eta_p^2 = .01$, and the two way *interaction* of group × time, $F(2, 131) = .43$, $p = .65$, $\eta_p^2 = .01$ were not significant (Table 1).

Motivation

To examine scores on the motivation scales, three 2 × 3 repeated measures analyses of variance (mixed ANOVA) were conducted for the three motivational tested sub-goals: *mastery goal, performance-approach goal* and *performance-avoidance goal*. In each analysis, group (Faded and Fixed) was treated as the between-subjects variable, and time (T1, T2, T3) as the within-subjects variable.

Mastery goal

The two way interaction of group x time was significant, $F(2, 131)=4.95$, $p<.01$, $\eta_p^2=.07$ (Table 2). Simple effect analyses comparing the three-time points for each study group separately indicated that while significant differences were found in the *Fixed* Group between the three time points, $F(2, 66)=5.83$, $p<.01$, $\eta_p^2=.15$, no significant differences were found between the three time points in the *Faded* Group, $F(2, 64)=.99$, $p=.38$, $\eta_p^2=.03$. Bonferroni analyses indicated that a *decrease* in scores between T1 and T3 ($d=-0.42$) was found in the *Fixed* Group ($p<.05$), with no significant differences in scores between T1 and T2 ($p=.26$) and between T2 and T3 ($p=.16$). Comparing the two study groups at each time point reveals that while no significant differences were found between the two groups in T1 ($p=.08$) and T2 ($p=.85$), higher scores in the *Faded* Group, compared to the *Fixed* Group, were found in T3 ($p<.05$, $d=0.33$) (Fig. 5).

Performance-approach goal

The two way i*nteraction* of group × time was significant, $F(2, 131)=5.17$, $p<.01$, $\eta_p^2=.07$ (Table 2). Simple effect analyses comparing the three-time points for each study group separately indicated no significant differences between the three time points in both groups, $F(2, 64)=2.34$, $p=.11$, $\eta_p^2=.07$ for the *Faded* Group and $F(2, 66)=2.86$, $p=.07$, $\eta_p^2=.08$ for the *Fixed* Group. Comparing the two study groups at each time point reveals that while no significant differences were found between the two groups in T1 ($p=.36$) and T2

Table 2 Means (and SD) and *F*-score of the motivation variables by the study groups fixed (n = 68) and faded (n = 66) and three-time points

Dependent variables	Study groups	Pre (T1)		Post (T2)		Long-term (T3)		F-Score	
		M	SD	M	SD	M	SD	F Time	F Interaction
Mastery goal1	Fixed	4.27	0.70	4.09	0.84	3.93	0.86	.36	4.95**
	Faded	4.02	0.95	4.12	0.91	4.21	0.89		
Performance-approach goal[a]	Fixed	3.33	1.29	3.25	1.24	3.03	1.26	.25	5.17**
	Faded	3.53	1.24	3.46	1.19	3.72	1.21		
Performance-avoidance goal[b]	Fixed	3.17	1.05	3.30	1.20	2.98	1.26	.80	7.59***
	Faded	3.23	1.09	3.31	1.12	3.67	1.15		

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

[a]Five-point Likert scale ranging from "1": Never, to "5": Always

[b]The measure of Performance-avoidance goal is inverted; thus, higher scores indicate less avoidance

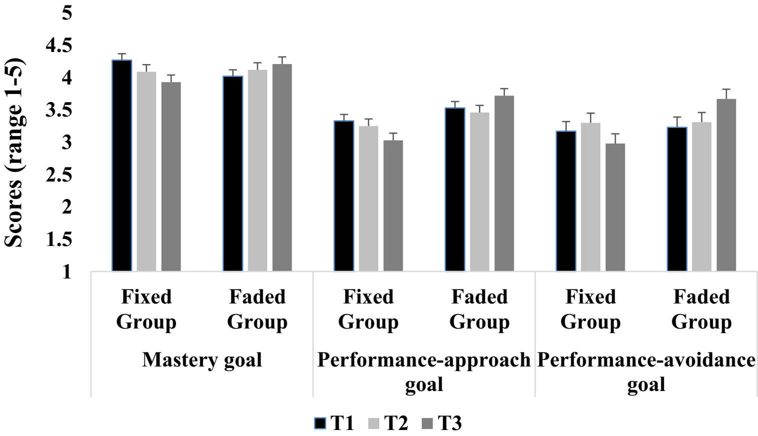


Fig. 5 Motivation scores of two study groups for three-time points

($p = .33$), higher scores in the *Faded* Group compared to the *Fixed* Group were found in T3 ($p < .01$, $d = 0.56$) (Fig. 5).

Performance-avoidance goal (scale inverted)

The two way *interaction* of group x time was significant, $F(2, 131) = 7.59$, $p < .001$, $\eta_p^2 = .10$ (Table 2). Simple effect analyses comparing the three time points for each study group separately indicated significant differences between the three time points in both groups, [$F(2, 64) = 5.72$, $p < .01$, $\eta_p^2 = .15$ for the *Faded* Group and $F(2, 66) = 3.20$, $p < .05$, $\eta_p^2 = .10$ for the *Fixed* Group]. Bonferroni analyses indicated that the *Faded* Group manifested *less avoidance* in T3 compared to T1($d = 0.34$) and T2 ($ps < .05$; $d = 0.33$), with no significant differences in scores between T1 and T2 ($p = .99$). However, the *Fixed* Group manifested significant *high*-avoidance in T3 compared to T2 ($d = -.34$), with no significant differences between T1 and T2 ($p = .95$) and between T1 and T3 ($p = .66$). Comparing the two study groups at each time point reveals that while no significant differences were found between the two groups in T1 ($p = .74$) and T2 ($p = .95$), higher scores in the *Faded* Group, compared to the *Fixed* Group, were found in T3 ($p < .001$, $d = 0.57$) (Fig. 5).

Sense making of problem solving

To examine the scores in the sense making of problem solving, a 2x3 repeated measures analysis of variance (mixed ANOVA) was conducted, with group (Faded and Fixed) as the between-subjects variable, and time (T1, T2, T3) as the within-subjects variable. The two way *interaction* of group x time was significant, $F(2, 131) = 3.04$, $p < .05$, $\eta_p^2 = .05$ (Table 1). Simple effect analyses comparing the three time points for each study group separately indicated significant differences between the three time points in both groups, [$F(2, 64) = 16.54$, $p < .001$, $\eta_p^2 = .34$ for the *Faded* Group and $F(2, 66) = 6.76$, $p < .01$, $\eta_p^2 = .17$ for the *Fixed* Group]. Bonferroni analyses indicated that while higher scores were found in T3 compared to T1 $d = 0.60$) and T2 $d = 0.68$) in the *Faded* Group ($ps < .001$), in the *Fixed* Group, higher scores were found in T3 compared to T2 ($p < .001$ $d = 0.45$), but not to T1 ($p = .36$). No significant differences were found in either group between T1 and T2 ($p = .99$)

for the *Faded* Group and $p = .21$ for the *Fixed* Group. Comparing the two study groups at each time point reveals that while no significant differences were found between the two groups in T1 *(*$p = .55$) and T2 ($p = .48$), higher scores in the *Faded* Group compared to the *Fixed* Group were found in T3 ($p < .05$, $d = 0.40$) (Fig. 6).

Discussion

The major finding by this study on *Fixed versus Faded* scaffolds on SRL and sense making benefit for the two metacognitive IMPROVE training groups, was that the faded scaffolds were more effective compared to fixed scaffolds on the tested variables, particularly on *lasting effect* (T3). These findings indicate that the opportunity to match the scaffolds autonomously to each SRL phase in the faded group was more effective than the static use of those scaffolds by the fixed group (Narciss et al. 2007). We discuss further the main findings for each variable.

Metacognition: knowledge and regulation of cognition

Differences between the groups were found particularly in the *knowledge* of cognition long-term awareness measure (T3), While the *Faded* Group exhibited a *higher* significant *increase*, the *Fixed* Group exhibited a significant *decrease* compared to the short-term effect (T2). But not for the *regulation* of cognition measure, whereas neither group manifested a statistically significant increase across the three-time intervals (T1, T2, T3). These findings raise two questions -

Q1 Why didn't the *Fixed* Group benefit from the short/lasting effect relating to the metacognition measures –knowledge and regulation aspects?

This result is surprising in light of previous studies in mathematics and in other knowledge fields which indicated that *fixed* metacognitive self-questions scaffolds of the *What, When, How*, and *Why* improve metacognition by reducing cognitive load, thus helping to increase WM capacity (Ariës et al. 2015; Kramarski and Fridman 2014). These contradictory findings can be explained in light of the different ages of participants in those studies.

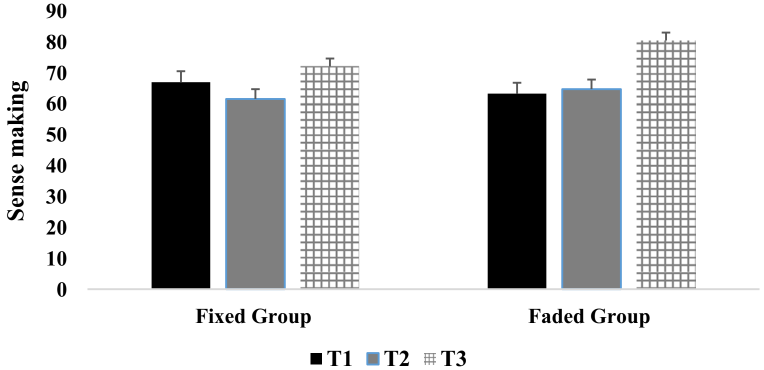


Fig. 6 Sense making of two study groups for three-time points

Whereas previous studies were mainly implemented with secondary/high school students and adults, the current study focuses on younger ages (9–10). Indeed, at that stage students begin to be mature enough to achieve initial awareness of metacognition, but the brief intervention (7 sessions) and the intensive continuous use of metacognitive self-questions in the *Fixed* Group might be too loading for the young students' knowledge attainment. As an initial stage of metacognition, lack of the knowledge aspect may further prevent the regulation of cognition aspect (Schraw 1998; Veenman et al. 2004).

Q2 Why did the *Faded* Group benefit significantly only on the knowledge of cognition (short/long-term) level, but not on the *regulation of cognition*? It seems that the autonomous learning that the *Faded* scaffold provided for these young students was beneficial for the first level of metacognition development (i.e., domain-specific knowledge focus), that is characterized as a lower path of thinking (Schraw 1998), but was not effective enough in short intervention to cope with the higher order level of metacognition (i.e., domain-general regulation focus).

Some research findings in mathematics emphasize the effect of visualized thinking strategies for assimilation of metacognitive problem-solving phases (Abdullah et al. 2014; Cleary et al. 2017; Rittle-Johnson and Star 2007). The lack of regulation aspect attainment could be explained by the linear visual strategy thinking that the prompting model raised (Fig. 1), that succeeded in building a linear mental model of knowledge awareness in comprehension (what?), strategy selection (how?) and initial reflection ability (why?) (Schraw 1998), but missed understanding of the regulation loop. Perhaps a systematic *visual*, *cyclical* practice of the SRL phases could help in internalizing that loop (Cleary et al. 2017; Pintrich et al. 2000; Zimmerman 2008). Future studies should further investigate the *Fixed versus Faded* scaffold effects with the two visualized scaffolds to prompt SRL, by comparing the groups once with the cyclical emphasis and then with the linear emphasis.

Confidence judgment

Q3 Why did both scaffold groups over-estimate their performance on T2 and then decrease on T3? In light of the importance of process measures to assess SRL in real time events (Azevedo 2014; Greene and Azevedo 2010), the calibration of confidence judgments' measure adds authentic data to students' metacognitive awareness of their performance on sense making, that adds understanding of the self-report metacognitive measures (knowledge and regulation). Interestingly, while students from both scaffolds groups manifested over-estimation of self-accuracy at the beginning of the study (T1) that further increased significantly at the end of the study (T2; short—term effect), they later became more realistic on the lasting effect (T3).

The dramatic confidence judgments' bias increase at the end of both scaffolds periods (T2) is explained in the literature as overwhelming from a new scaffold experience that raised young student's self-efficacy feelings of success in solving the task (e.g., Pintrich 2000). However, the realistic self-estimation on T3 can be explained as a result of internalizing the monitoring and reflection mechanism they were trained to realize accurately their own success. This conclusion is supported by findings of previous studies, according to which practicing confidence judgments' accuracy while teaching a content field improves the ability in this type of judgement among nine-year-old students (Gidalevich and Kramarski

2017; Roderer and Roebers 2010). They also extend the findings by Zimmerman and colleagues (2011), who found that confidence judgments' ability among college students was improved during a longer practice (15 weeks) compared to our fourth graders (7 weeks). This finding has practical implications for young students' achievement improvement. Because a learner with high confidence judgment is unlikely to go back to recheck and correct his mistakes, as such he may mistakenly fall into an at-risk category as a low achiever (Labuhn et al. 2010; Shin et al. 2007).

Motivation: mastery goal, performance-approach goal, performance- avoidance goal

Q4 Why did the Fixed Group *decrease* on mastery goal measure compared to the Faded Group on the lasting effect (T3)?, while the Faded Group manifested *higher* mastery goal, performance-approach goals and *less* performance-avoidance goals in learning mathematics compared to the Fixed Group. This finding supports the underpinning rationale of this study, that autonomous engagement in learning with the faded scaffolds raises students' motivational goal-orientation (Adler et al. 2016).

Furthermore, the motivational achievements can be explained regarding the groups' metacognition gains. It seems that in the Faded Group the positive trend on the three kinds of goals, mastery, performance-approach and avoidance is a result of the metacognitive knowledge and initial self-regulation improvement (Tables 1, 2).

This is in line with theoretical assumptions about relations between metacognition and motivation (Pintrich 2000). Metacognition is regarded as a superordinate ability to direct and regulate cognitive, motivational and behavioral processes in order to achieve specific goals (Ifenthaler 2012; Pintrich 2000).

The similar findings trend in metacognition (knowledge aspect) and motivation support the assertion that SRL dimensions and processes are linked (Kim and Pekrun 2014; Pintrich et al. 2000; Schunk 1991) and may be influenced by the same specific instructional intervention's prompts (Ifenthaler 2012). This conclusion is supported also by the correlations between metacognition and motivation in Tables 6 and 7 in Appendix C.

Q5 Why didn't the Faded Group increase significantly their mastery goal along the time intervals? The lack of mastery goal achievements increase can be explained by the nature of the *indirect* scaffold on motivation that hindered the effect on the mastery goal achievement that as such is conceived of as a transfer skill (Kistner et al. 2010). Future studies should further compare the effects of two faded scaffold groups: one directed explicitly to motivation and the other directed explicitly to metacognition as implemented in the study, to find which of the indirect SRL aspects, metacognition or motivation, is more affected by the faded scaffolds.

The SRL findings (metacognition, CJ, motivation) add understanding with regard to the role of the faded scaffolds in young students in fostering SRL in mathematics as a proactive process that does not merely happen *to* students, but rather occurs autonomously within themselves (Zimmerman 2008). It expands previous findings on IMPROVE self-questioning prompts as a fixed scaffold with young students (e.g., fifth grade; Tzohar-Rozen and Kramarski 2014), that focused only on SRL effect at the end of the program, without allowing time for assimilation and preservation of the program and testing its long-term effect, as in the current study.

Sense making of problem solving: short term and long-term effect

Q6 What makes the difference between the two scaffolds types: Fixed versus Faded on the tested variables? Findings indicated that *both* groups *increased* their sense making only in the *long-term* effect (T3) compared to the *immediate* effect (T2), whereas a greater increase was found in the Faded Group ($d=0.68$) compared to the Fixed Group ($d=0.45$). This is in line with research showing a larger impact of metacognitive effects in follow-up learning sessions than earlier measures of short-term effects (Bannert et al. 2015).

These findings support previous conclusions, that metacognitive scaffolds can serve as a learning strategy that enhances long-term memorization and re-use of these learning activities to develop awareness about the relevance of learning activities (e.g., Ariës et al. 2015; Goldberg 2010; Kuhn and Dean 2004;). The differences between the two scaffold types (Fixed vs. Faded) on the sense making of problem solving can be explained by relating to the way that the information processing in each group is stimulated. The fixed scaffolds enabled learners to walk through SRL problem solving phases with explicit SRL question prompts step by step, thereby emphasizing local (concrete) processes guided by *less* autonomy in self-selected prompts that should be adapted to the SRL phases. This naturally leads to low-road transfer (Salomon and Perkins 1989) to the sense making of problem-solving long-term retention (T3). In contrast, the Faded graduated model offered more scaffolding to abstract information processing by students' exposure to only one example of full problem-solving prompting phases. This was followed by isomorphic tasks to be solved with self-complementing the withdrawn prompts matched to the relevant phases in the problem-solving process. Such autonomous mindful abstraction establishes the high road (generalization) to far transfer (T3; Salomon and Perkins 1989).

Our findings extend earlier findings on faded scaffolds that were achieved mainly in the science domain but did not lead to conclusive results (e.g., Bulu and Pedersen 2010). The Faded Group's sense making gains seems to be a result of the positive effects that emerged in the SRL (metacognition and motivation) measures, showing the *synergic effect* of scaffolding explicitly and directly on one aspect of the SRL model as metacognition awareness and its indirect effect on related aspects of motivation. This finding is most important, because if students are not motivated to engage in a particular task, they are unlikely to be strategic in that process which assists students to achieve a deeper understanding (Efklides 2011).

Conclusions, implications, limitations and future research

This study makes an important contribution to the literature on young students' (9–10), indicating that the fading effect with the graduated model is more significant compared to the fixed effect, particularly in the lasting effect. Theoretically, the current findings add evidence to the two scaffold type models (Fixed vs. Faded), and transfer issues (long-term effect) that offer merits in advancing our understanding about scaffolding theory. Current scaffolding literature does not clearly indicate what kinds o scaffolds to maintain and what kinds of scaffolds to fade at a certain point in time (Belland et al. 2017; Bulu and Pedersen 2010).

Methodologically, the study contributes findings on a set of SRL self-report measurements (metacognition, motivation), authentic self-assessment (calibration of CJ), and performance in a non-prompted sense-making task at the end of the program (short-term effect), and three-months later to assess lasting effect. Practically, the study offered two programs based on self-question models (Fixed, Faded), to guide acquisition, activation, and application to proactively promote metacognition of young students, and are adaptable to different ages, various mathematical topics and to other learning domains. This is in line with recent studies asserting that metacognition should begin from the primary school level (Kistner et al. 2010).

Despite this study's potential contributions, several of its limitations deserve consideration. In this study SRL was measured via self-reports in two classes of each group that indicated relatively small/moderate significant effect sizes, that might express a first type error increase. Further investigation is suggested with a larger sample in order to determine the fading contribution to scaffolding programs.

Indeed, findings indicated that the two groups were homogeneous in relating to the measures implemented in the study (see Appendix A). However, there might be additional measures that were not tested and could affect the homogeneity of the groups, such as norms of class talk and verbal abilities that can affect problem solving performance. Future studies should also incorporate class background measures in a follow-up study.

In light of the importance of involving *trace* methodology in SRL studies (e.g., Zimmerman 2008) an additional measure of metacognition improvement is suggested, as students' thinking aloud while solving a task or videotaped discourse in class to shed light on sense making. This may help triangulate the testing variables and uncover students' sequential and temporal metacognitive and motivational aspects exhibited during various solving activities, by focusing on *when* each aspect appeared and by *whom* over the timeline of the learning activity, its duration and *how* students act (Azevedo 2014; Winne and Perry 2000). Finally, an additional measure to shed light on brain working processes such as cognitive load and WM while occurring in the short term versus long-term effect is suggested to assess the two study groups (Ariës et al. 2015; Goldberg 2010).

In line with research showing a larger impact of metacognitive effects in follow-up learning sessions than measured earlier in short-term effects (Bannert et al. 2015), we recommend that this study be replicated by focusing on a longer duration of process assessment with a larger sample, while testing immediately after the training as well as following up on two time lines of lasting effects, to exhibit how students internalize and practice metacognitive habits without training (Kramarski et al. 2013).

Future studies should examine additional kinds of fading, conditions, domains, SRL components (i.e., motivation) and ages. In the current study the faded progression was based on the SRL metacognition prompts oriented to the three phases practiced by a set of 4 tasks in each session, starting with a fixed prompting example and continuing with the graduated faded phases. It might be that this approach is confounded by student's content experience. To face this limitation, we suggest that fading scaffolds could also be tested if the intensity of removal is not fixed, but adaptable to the individual needs of the learner (Cabello and Sommer Lohrmann 2018).

Furthermore, in the current study the faded sequence of SRL phases oriented to the task solution started with complementing the reflection prompt that is more general and abstract than the planning prompt, that is specific and more concrete. In future studies it is suggested to test an inverse fading sequence by starting with the given abstract prompting phase (reflection) to complementing the missing concrete prompting phase. Finally, studies should also focus on the implications of the metacognitive scaffolds (fixed/faded) for different students' ages and profiles (e.g., SRL, achievements), to learn how to adapt the models to learners with *diverse* learning needs, because novices might need a fuller set of scaffolds for a longer period to support their problem-solving performance (Tomlinson 2005).

To conclude , the study is unique in comparing the value of the fixed versus graduated fading scaffold directed to SRL and not to domain knowledge. It supports theoretical claims that fading scaffolds have a greater effect on students' autonomous SRL (metacognition, motivation) and improvements in sense making, particularly on the long-term retention effect (Ge et al. 2012; Puntambekar and Hubscher 2005).

Acknowledgements We confirm that we have reported all measures, conditions, data exclusions, and how we determined our sample sizes. This research was supported by Oranim Academic College of Education.

Compliance with Ethical Standards

Conflict of interest The authors declare that they have no conflict of interest.

Appendix A

See Appendix Tables 3 and 4.

Table 3 Means (and SD) of the dependent variables by class at each time point in the Fixed group

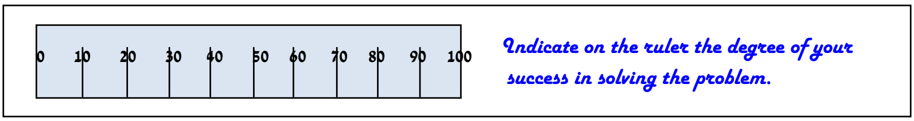
		Class 1 (n = 35)		Class 2 (n = 33)				
Dependent variables	Time	M	SD	M	SD	t	p	Cohen's d
Knowledge of cognition	T1	4.01	0.56	4.01	0.63	− .04	.97	0.00
	T2	4.19	0.56	4.11	0.63	.61	.54	0.13
	T3	4.00	0.63	3.92	0.64	.57	.57	0.13
Regulation of cognition	T1	3.77	0.62	3.87	0.59	− .72	.48	0.17
	T2	3.90	0.68	3.96	0.62	− .40	.69	0.09
	T3	3.85	0.62	3.77	0.71	.47	.64	0.12
Confidence Judgements	T1	− 10.21	39.60	− 19.09	38.02	.94	.35	0.23
	T2	− 45.57	34.44	− 38.58	39.28	− .78	.44	0.19
	T3	− 3.29	28.59	− 5.76	29.88	.35	.73	0.08
Mastery goal	T1	4.21	0.76	4.33	0.63	− .73	.47	0.17
	T2	4.02	0.95	4.17	0.70	− .75	.46	0.18
	T3	3.93	0.90	3.93	0.84	− .00	.99	0.00
Performance-approach goal	T1	3.30	1.31	3.35	1.28	− .14	.89	0.04
	T2	3.25	1.31	3.26	1.18	− .03	.97	0.01
	T3	3.10	1.23	2.96	1.30	.44	.66	0.11
Performance-avoidance goal	T1	3.14	1.08	3.20	1.03	− .25	.80	0.06
	T2	3.39	1.25	3.20	1.16	.64	.52	0.16
	T3	3.00	1.25	2.96	1.30	.15	.88	0.03
Sense making of problem solving	T1	62.38	26.84	71.72	28.94	− 1.38	.17	0.33
	T2	59.05	25.03	64.09	28.26	− .78	.44	0.19
	T3	71.68	22.22	72.35	24.16	− .12	.91	0.03

Table 4 Means (and SD) of the dependent variables by class at each time point in the Faded group

		Class 1 (n = 33)		Class 2 (n = 33)				
Dependent variables	Time	M	SD	M	SD	t	p	Cohen's d
Knowledge of cognition	T1	4.04	0.71	3.90	0.84	.73	.47	0.18
	T2	4.28	0.58	4.03	0.77	1.48	.14	0.37
	T3	4.37	0.56	4.09	0.69	1.78	.08	0.45
Regulation of cognition	T1	3.79	0.71	3.84	0.86	− .25	.81	0.06
	T2	4.06	0.63	3.81	0.74	1.52	.13	0.36
	T3	4.16	0.75	3.79	0.88	1.83	.07	0.45
Confidence Judgements	T1	− 11.21	24.14	− 17.27	37.48	.78	.44	0.19
	T2	− 32.27	44.76	− 41.21	38.55	.87	.39	0.21
	T3	7.35	22.33	− 1.36	27.14	1.42	.16	0.24
Mastery goal	T1	3.95	0.83	4.09	1.07	− .58	.56	0.15
	T2	4.07	0.85	4.17	0.97	− .45	.65	0.11
	T3	4.30	0.76	4.13	1.01	.80	.42	0.19
Performance-approach goal	T1	3.78	1.01	3.27	1.40	1.69	.10	0.42
	T2	3.67	1.02	3.25	1.32	1.44	.15	0.36
	T3	3.92	0.92	3.52	1.43	1.37	.18	0.33
Performance-avoidance goal	T1	3.44	0.98	3.02	1.17	1.58	.12	0.39
	T2	3.48	0.95	3.15	1.25	1.22	.23	0.30
	T3	3.80	0.90	3.54	1.35	.92	.36	0.23
Sense making of problem solving	T1	70.20	27.95	57.58	30.57	1.75	.09	0.43
	T2	68.18	24.43	61.11	25.23	1.16	.25	0.28
	T3	82.20	17.41	78.60	19.45	.78	.43	0.19

Appendix B

1. Confidence Judgments—CJ: Assessing the correctness of one's performance after the solution.



2. Sense making of problem solving task

Fourth grade students went on a kayak trip. There are small kayaks for 4 students and large kayaks for 6 students. Kayaks only leave when they are full. How many students were on the trip, if 12 large kayaks and 8 small kayaks left for the trip and two students had to wait for additional students to go on the kayaks?

Present the solution path:

Answer: ____________ students.

Expected answer	Grading
106 students Path of the solution: $12 \times 6+8 \times 4+2=106$ Or according to the phases: 12 X 6 = 72 8 X 4 = 32 72 + 32 + 2 = 106	100 pts—correct answer and correct presentation of the solution path 50 pts—correct presentation of the solution path but wrong answer, or correct final answer without presentation of the solution path or a wrong presentation of the solution path 0 pts—wrong answer without presenting the solution path or with presenting a wrong solution path

Appendix C: Pearson correlation coefficients between the three aspects of the SRL (N = 134)

See Appendix Tables 5, 6 and 7.

Table 5 T1

Dependent variables	1	2	3	4	5	6	7
Knowledge of cognition (1)	1	.76***	− .02	.61***	.37***	.47***	.23**
Regulation of cognition (2)		1	− .08	.65***	.43***	.45***	.08
Confidence Judgements (3)			1	.03	− .12	− .16	.28**
Mastery goal (4)				1	.25**	.28***	.16
Performance-approach goal (5)					1	.62***	.02
Performance-avoidance goal (6)						1	.04
Sense making of problem solving (7)							1

Table 6 T2

Dependent variables	1	2	3	4	5	6	7
Knowledge of cognition (1)	1	.75***	.03	.47***	.31***	.30***	.14
Regulation of cognition (2)		1	− .05	.48***	.43***	.41***	.02
Confidence Judgements (3)			1	− .01	− .02	.01	.14
Mastery goal (4)				1	.26**	.24**	.11
Performance-approach goal (5)					1	.69***	.02
Performance-avoidance goal (6)						1	− .04
Sense making of problem solving (7)							1

Table 7 T3

Dependent variables	1	2	3	4	5	6	7
Knowledge of cognition (1)	1	.78***	− .06	.47***	.49***	.43***	.22*
Regulation of cognition (2)		1	− .02	.62***	.50***	.48***	.14
Confidence Judgements (3)			1	− .08	.04	.09	.08
Mastery goal (4)				1	.39***	.35***	.25**
Performance-approach goal (5)					1	.82***	− .02
Performance-avoidance goal (6)						1	− .04
Sense making of problem solving (7)							1

*p < .05, **p < .01, ***p < .001

References

Abdullah, N., Halim, L., & Zakaria, E. (2014). VStops: A thinking strategy and visual representation approach in mathematical word problem solving toward enhancing STEM literacy. *Eurasia Journal of Mathematics, Science & Technology Education, 10*(3), 165–174.

Adler, I., Schwartz, L., Madjar, N., & Zion, M. (2016). Reading between the lines: The effect of contextual factors on student motivation throughout an open inquiry process. *Science Education, 102*(4), 820–855.

Al-Harthy, I., & Was, C. (2010). Goals, efficacy and metacognitive self-regulation: A path analysis. *International Journal of Education, 2*(1), 1–2.

Ames, C. (1992). Classrooms: Goals, structures, and student motivation. *Journal of Educational Psychology, 84*(3), 261–271.

Ariës, R. J., Ghysels, J., Groot, W., & Van den Brink, H. M. (2015). Is working memory training effective in enhancing school based reasoning achievements? A systematic review. TIER working paper series.

Ariës, R. J., Groot, W., & Van den Brink, H. M. (2014). Improving reasoning skills in secondary history education by working memory training. *British Educational Research Journal, 41*(2), 210–228. https://doi.org/10.1002/berj.3142.

Azevedo, R. (2014). Issues in dealing with sequential and temporal characteristics of elf- and socially-regulated learning. *Metacognition and Learning, 9*(2), 217–228. https://doi.org/10.1007/s11409-014-9123-1.

Baddeley, A. (1992). Working memory. *Science, 255*(5044), 556–559.

Bannert, M., & Mengelkamp, C. (2013). Scaffolding hypermedia learning through metacognitive prompts. In R. Azevedo & V. Aleven (Eds.), *International handbook of metacognition and learning technologies* (pp. 171–186). New York: Springer. https://doi.org/10.1007/978-1-4419-5546-3_12.

Bannert, M., & Reimann, P. (2012). Supporting self-regulated hypermedia learning through prompts. *Instructional Science, 40*(1), 193–211. https://doi.org/10.1007/s11251-011-9167-4.

Bannert, M., Sonnenberg, C., Mengelkamp, C., & Pieger, E. (2015). Short- and long-term effects of students' self-directed metacognitive prompts on navigation behavior and learning performance. *Computers in Human Behavior, 52,* 293–306.

Belland, B. R., Walker, A. E., Kim, N. J., & Lefler, M. (2017). Synthesizing results from empirical research on computer-based scaffolding in STEM education: A Meta-Analysis. *Review of Educational Research, 87*(2), 309–344. https://doi.org/10.3102/0034654316670999.

Boekaerts, M., & Corno, L. (2005). Self-regulation in the classroom: A perspective on assessment and intervention. *Applied Psychology: An International Review, 54,* 199–231.

Brown, A. (1987). Metacognition, executive control, self-regulation, and other more mysterious mechanisms. In F. Weinert & R. Kluwe (Eds.), *Metacognition, motivation and understanding* (pp. 65–116). Hillsdale, NJ: Erlbaum.

Bulu, S., & Pedersen, S. (2010). Scaffolding middle school students' content knowledge and ill-structured problem solving in a problem-based hypermedia-learning environment. *Educational Technology Research and Development, 58,* 507–529. https://doi.org/10.1007/s11423-010-9150-9.

Cabello, V. M., & Sommer Lohrmann, M. E. (2018). In T. Andre (Ed.), *Advances in human factors in training, education, and learning sciences, advances in intelligent systems and computing 596* (pp. 350–360). https://doi.org/10.1007/978-3-319-60018-5_34.

Chatzistamatiou, M., Dermitzaki, I., Efklides, A., & Angeliki, L. (2015). Motivational and affective determinants of self-regulatory strategy use in elementary school mathematics. *Educational Psychology, 35*(7), 835–850. https://doi.org/10.1080/01443410.2013.822960.

Cleary, T. J., Velardi, B., & Schnaidman, B. (2017). Effects of the self-regulation empowerment program (SREP) on middle school students' strategic skills, self-efficacy, and mathematics achievement. *Journal of School Psychology, 64,* 28–42. https://doi.org/10.1016/j.jsp.2017.04.004.

Davis, E. A. (2003). Prompting middle school science students for productive reflection: Generic and directed prompts. *Journal of the Learning Sciences*, *12*(1), 91–142.

Devoldr, A., van Braak, J., & Tondeur, J. (2012). Supporting self-regulated learning in computer-based learning environments: systematic review of effects of scaffolding in the domain of science education. *Journal of Computer Assisted Learning, 28*(6), 557–573.

Dignath, C., Buettner, G., & Langfeldt, H. P. (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*(2), 101–129.

Dweck, C. S. (1986). Motivational processes affecting learning. *American Psychologist, 41,* 1040–1048.

Dweck, C. S., & Leggett, E. L. (1988). A social-cognitive approach to motivation and personality. *Psychological Review, 95,* 256–272.

Efklides, A. (2011). Interactions of metacognition with motivation and affect in self-regulated learning: The MASRL Model. *Educational Psychologist, 46*(1), 6–25.

Elliot, A. (1997). Integrating the ''classic'' and ''contemporary'' approaches to achievement motivation: A hierarchical model of approach and avoidance achievement motivation. In M. L. Maehr & P. R. Pintrich (Eds.), *Advances in motivation and achievement* (Vol. 10, pp. 143–179). Greenwich, CT: JAI Press.

Flavell, J. H. (1979). Metacognition and cognitive monitoring: A new area of cognitive developmental inquiry. *American Psychologist, 34,* 906–911. https://doi.org/10.1037/0003-066X.34.10.906.

Ge, X., Law, V., & Huang, K. (2012). Diagnosis, supporting, and fading: A scaffolding design framework for adaptive e-learning systems. In H. Wang (Ed.), *Interactivity in E-learning: case studies and frameworks* (pp. 116–142). Hershey, PA: IGI Global.

Gidalevich, S., & Kramarski, B. (2017). Metacognitive guidance for self-regulation judgements in various phases: A thinking aloud analysis in mathematics. *Hellenic Journal of Psychology*, *14*, 88–113.

Goldberg, E. (2010). *Het sturende brein: Onze hersenen in een complexe wereld [The new executive brain: Frontal lobes in a complex world].* Amsterdam: Wereldbibliotheek.

Greene, J. A., & Azevedo, R. (2010). The measurement of learners' self-regulated cognitive and metacognitive processes while using computer-based learning environments. *Educational Psychologist, 45*(4), 203–209. https://doi.org/10.1080/00461520.2010.515935.

Hoffman, B., & Spatariu, A. (2008). The influence of self-efficacy and metacognitive prompting on math problem-solving efficiency. *Contemporary Educational Psychology, 33*(4), 875–893.

Huff, J. D., & Nietfeld, J. L. (2009). Using strategy instruction and confidence judgments to improve metacognitive monitoring. *Metacognition Learning, 4*(2), 161–176. https://doi.org/10.1007/s11409-009-9042-8.

Ifenthaler, D. (2012). Determining the effectiveness of prompts for self-regulated earning in problem-solving scenarios. *Educational Technology & Society, 15*(1), 38–52.
Israeli Ministry of Education. (2005). *School efficiency and growth measures test for 4PthP graders—Version A*. Jerusalem: Department of Assessment and Measurement. **in Education**.
Jaakkola, T., & Veermans, K. (2018). Exploring the effects of concreteness fading across grades in elementary school science education. *Instructional Science, 46*(2), 185–207.
Kaplan, A., & Maehr, L. M. (2002). Adolescents' achievement goals: situating motivation in socio-cultural contexts. In T. Urdan & F. Pajaers (Eds.), *Academic motivation of adolescents. Adolescence and education* (Vol. 2, pp. 125–167). Greenwich, CT: Iinformation Age.
Kester, L., & Kirschner, P. A. (2009). Effects of fading support on hypertext navigation and performance in student-centered e-learning environments. *Interactive Learning Environments, 17*(2), 165–179. https://doi.org/10.1080/10494820802054992.
Kim, C., & Pekrun, R. (2014). Emotions and motivation in learning and performance. In J. M. Spector, M. D. Merrill, J. Elen, & M. J. Bishop (Eds.), *Handbook of research on educational communications and technology* (pp. 65–75). New York: Springer.
Kistner, S., Rakoczy, K., Otto, B., Dignath-van Ewijk, C., Büttner, G., & Klieme, E. (2010). Promotion of self-regulated learning in classrooms: Investigating fre-quency, quality, and consequences for student performance. *Metacognition and Learning, 5*(2), 157–171.
Koedinger, K. R., & Aleven, V. (2007). Exploring the assistance dilemma in experiments with cognitive tutors. *Educational Psychology Review, 19,* 239–264.
Kramarski, B. (2017). Teachers as agents in promoting students' SRL and performance: Applications for teachers' dual-role training program. In D. H. Schunk & J.A. Greene (Eds.), *Handbook of self-regulation of learning and performance* (pp. 223–239). New York: Springer.
Kramarski, B., & Fridman, S. (2014). Solicited versus unsolicited metacognitive prompts for fostering mathematical problem-solving using multimedia. *Journal of Educational Computing Research*, *50*(3), 285–314.
Kramarski, B., & Mevarech, Z. R. (2003). Enhancing mathematical reasoning in the classroom: Effects of cooperative learning and metacognitive training. *American Educational Research Journal*, *40*, 281–310. https://doi.org/10.3102/00028312040001281.
Kramarski, B., & Revach, T. (2009). The challenge of self-regulated learning in mathematics teachers' professional training. *Educational Studies in Mathematics*, *72*(3), 379–399.
Kramarski, B., Weiss, I., & Sharon, S. (2013). Generic versus context-specific prompts for supporting self-regulation in mathematical problem solving among students with low or high prior knowledge. *Journal of Cognitive Education and Psychology*, *12*, 197–214. https://doi.org/10.1891/1945-8959.12.2.97.
Kuhn, D., & Dean, D. (2004). A bridge between cognitive psychology and educational practice. *Theory into Practice, 43*(4), 268–273.
Labuhn, A. S., Zimmerman, B. J., & Hasselhorn, M. (2010). Enhancing students' self-regulation and mathematics performance: the influence of feedback and self-evaluative standards. *Metacognition and Learning, 5*(2), 173–194. https://doi.org/10.1007/s11409-010-9056-2.
Lee, H. & Songer, N. B. (2004). Expanding an understanding of scaffolding theory using an inquiry-fostering science program. Understanding scafolding. *The Journal of the Learning Sciences, October*.
Mayer, R. E. (2009). *Multimedia learning* (2nd ed.). New York: Cambridge University Press.
McNeill, K. L., Lizotte, D. J., Krajcik, J., & Marx, R. W. (2006). Supporting students' construction of scientific explanations by fading scaffolds in instructional materials. *Journal of the Learning Sciences, 15*(2), 153–191.
Meece, J. L. (1994). Individual and classroom differences in students' achievement goal patterns. In D. Schunk & B. Zimmerman (Eds.), *Self-regulated learning: Definitions and issues* (pp. 25–44). Hillsdale, NJ: Lawrence Erlbaum Associates.
Mevarech, Z. R., & Kramarski, B. (1997). IMPROVE: A multidimensional method for teaching mathematics in heterogeneous classroom. *American Educational Research Journal*, *34*, 365–395. https://doi.org/10.3102/00028312034002365.
Mevarech, Z. R., & Kramarski, B. (2014). *Critical maths for innovative societies: The role of metacognitive pedagogies*. Paris: OECD. https://doi.org/10.1787/9789264223561-en.
Midgley, C., Maehr, M., Hruda, L., Anderman, E., Anderman, L., Freeman, K., …Urdan, T. (2000). *Manual for the patterns of adaptive learning scales*. Ann Arbor, MI: University of Michigan.
Mihalca, L., Mengelkamp, C., & Schnotz, W. (2017). Accuracy of metacognitive judgments as a moderator of learner control effectiveness in problem-solving tasks. *Metacognition and Learning, 12*(3), 357–379. https://doi.org/10.1007/s11409-017-9173-2.

Moos, D. C., & Ringdal, A. (2012). Self-regulated learning in the classroom: A literature review on the teacher's role. *Education Research International*. https://doi.org/10.1155/2012/423284.

Müller, N. M., & Seufert, T. (2018). Effects of self-regulation prompts in hypermedia learning on learning performance and self-efficacy. *Learning and Instruction, 58,* 1–11. https://doi.org/10.1016/j.learninstruc.2018.04.011.

Narciss, S., Proske, A., & Koerndle, H. (2007). Promoting self-regulated learning in webbased learning environments. *Computers in Human Behavior, 23,* 1126–1144.

National Council of Teachers of Mathematics (NCTM). (2009). *Focus on HIGH SCHOOL MATHEMATICS: Reasoning and sense making*. Reston, VA: NCTM.

Nelson, T. O., & Narens, L. (1994). *Why investigate metacognition?*. Cambridge, MA: The MIT Press.

Nolen, S. B. (1988). Reasons for studying: Motivational orientation and strategies. *Cognition and Instruction, 5,* 269–287.

Panadero, E., Tapia, J. A., & Huertas, J. A. (2012). Rubrics and self-assessment scripts effects on self-regulation, learning and self-efficacy in secondary education. *Learning and Individual Differences, 22*(6), 806–813.

Patrick, H., Ryan, A. M., & Kaplan, A. (2007). Early adolescents' perceptions of the classroom social environment, motivational beliefs, and engagement. *Journal of Educational Psychology, 99,* 83–98.

Pea, R. D. (2004). The social and technological dimensions of scaffolding and related theoretical concepts for learning, education, and human activity. *The Journal of the Learning Sciences, 13*(3), 423–451.

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: Academic.

Pintrich, P. R. (2002). The role of metacognitive knowledge in learning, teaching, and assessing. *Theory Into Practice, 41*(4), 219–225.

Pintrich, P. R., Wolters, C., & Baxter, G. (2000). Assessing metacognition and self-regulated learning. In G. Schraw & J. Impara (Eds.), *Issues in the measurement of metacognition* (pp. 43–97). Lincoln, NE: Buros Institute of Mental Measurements.

Puntambekar, S., & Hubscher, R. (2005). Tools for scaffolding students in a complex learning environment: What have we gained and what have we missed? *Educational Psychologist, 40*(1), 1–12. https://doi.org/10.1207/s15326985ep4001_1.

Rittle-Johnson, B., & Star, J. R. (2007). Does comparing solution methods facilitate conceptual and procedural knowledge? An experimental study on learning to solve equations. *Journal of Educational Psychology, 99*(3), 561–574.

Roderer, T., & Roebers, C. M. (2010). Explicit and implicit confidence judgments and developmental differences in metamemory: An eye-tracking approach. *Metacognition and Learning, 5*(3), 229–250. https://doi.org/10.1007/s11409-010-9059-z.

Roebers, C. M., Krebs, S. S., & Roderer, T. (2014). Metacognitive monitoring and control in elementary school children: Their interrelations and their role for test performance. *Learning and Individual Differences, 29,* 141–149.

Salomon, G., & Perkins, D. N. (1989). Rocky roads to transfer: Rethinking mechanism of a neglected phenomenon. *Educational Psychologist, 24*(2), 113–142. https://doi.org/10.1207/s15326985ep2402_1.

Schoenfeld, A. H. (1992). Learning to think mathematically: Problem solving, metacognition, and sense making in mathematics. In D. A. Grouws (Ed.), *Handbook of research on mathematics teaching and learning* (pp. 165–197). New York: MacMillan.

Schraw, G. (1998). Promoting general metacognitive awareness. *Instructional Science, 26*(1–2), 113–125.

Schraw, G. (2009). A conceptual analysis of five measures of metacognitive monitoring. *Metacognition Learning, 4*(1), 33–45. https://doi.org/10.1007/s11409-008-9031-3.

Schraw, G., & Dennison, R. S. (1994). Assessing metacognitive awareness. *Contemporary Educational Psychology, 19,* 460–475.

Schunk, D. H. (1991). Self-efficacy and academic motivation. *Educational Psychologist, 26,* 207–232.

Seo, D., & Kim, J. (2001). Expanding a goal mediational model: The Korean elementary school math class. *Academic Exchange Quarterly, 5,* 177–183.

Shin, H., Bjorklund, D. F., & Beck, E. F. (2007). The adaptive nature of children's overestimation in a strategic memory task. *Cognitive Development, 22,* 197–212.

Sperling, R. A., Howard, B. C., Miller, L. A., & Murphy, C. (2002). Measures of children's knowledge and regulation of cognition. *Contemporary Educational Psychology, 27*(1), 51–79. https://doi.org/10.1006/ceps.2001.1091.

Sweller, J. (1988). Cognitive load during problem solving: effects on learning. *Cognition and Science, 12,* 257–285.

Tawfik, A. A., Law, V., Ge, X., Xing, W., & Kim, K. (2018). The effect of sustained versus faded scaffolding on students' argumentation in ill-structured problem solving. *Computers in Human Behavior*. https://doi.org/10.1016/j.chb.2018.01.035.

Tomlinson, C. A. (2005). Quality curriculum and instruction for highly able students. *Theory into Practice, 44*(2), 160–166.

Trends in International Mathematics and Science Study—TIMSS. (2011). Mathematics Frameworks, 2011. Retrieved from http://timss.bc.edu/index.html. Accessed 28 Nov 2018.

Tzohar-Rozen, M., & Kramarski, B. (2014). Metacognition, motivation and emotions: Contribution of self-regulated learning to solving mathematical problems. *Global Education Review*, *1*(4), 76–95.

Veenman, M. V. J., Van Hout-Wolters, B. H. A. M., & Afflerbach, P. (2006). Metacognition and learning: Conceptual and methodological considerations. *Metacognition and Learning, 1,* 3–14.

Veenman, M. V. J., Wilhelm, P., & Beishuizen, J. J. (2004). The relation between intellectual and metacognitive skills from a developmental perspective. *Learning and Instruction*, *14*(1), 89–109. https://doi.org/10.1016/j.learninstruc.2003.10.004.

Vrugt, A., & Oort, F. J. (2008). Metacognition, achievement goals, study strategies and academic achievement: Pathways to achievement. *Metacognition & Learning, 30,* 123–146.

Watts, T. W., Clements, D. H., Sarama, J., Wolfe, C. B., Spitler, M. E., & Bailey, D. H. (2016). Does early mathematics intervention change the processes underlying children's learning? *Journal of Research on Educational Effectiveness*. https://doi.org/10.1080/19345747.2016.1204640.

Weinstein, C. E., Acee, T. W., & Jung, J. (2011). Self-regulation and learning strategies. *New Directions for Teaching and Learning, 126,* 45–53.

Winne, P. H. (1996). A metacognitive view of individual differences in self-regulated learning. *Learning and Individual Differences, 8,* 327–353.

Winne, P., & Perry, N. (2000). Measuring self-regulated learning. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation* (pp. 531–566). San Diego, CA: Academic Press.

Wolters, C. A. (2004). Advancing achievement goal theory: Using goal structures and goal-orientations to predict students' motivation, cognition and achievement. *Journal of Educational Psychology, 96,* 236–250.

Zheng, L. (2016). The effectiveness of self-regulated learning scaffolds on academic performance in computer-based learning environments: a meta-analysis. *Asia Pacific Education Review, 17*(2), 187–202.

Zimmerman, B. J. (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: Academic Press.

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

Zimmerman, B. J., Moylan, A., Hudesman, J., White, N., & Flugman, B. (2011). Enhancing self-reflection and mathematics achievement of at-risk urban technical college students. *Psychological Test and Assessment Modeling, 53*(1), 108–127.

Zusho, A., & Edwards, K. (2011). Self-regulation and achievement goals in the college classroom. *New Directions for Teaching and Learning, 126,* 117–124.