Educational Research Review 28 (2019) 100292

Contents lists available at ScienceDirect

Educational Research Review

journal homepage: www.elsevier.com/locate/edurev

Self-regulated learning partially mediates the effect of self-regulated learning interventions on achievement in higher education: A meta-analysis

Renée S. Jansen[a,*], Anouschka van Leeuwen[a], Jeroen Janssen[a], Suzanne Jak[b], Liesbeth Kester[a]

[a] Department of Education, Utrecht University, the Netherlands
[b] Methods and Statistics, Research Institute of Child Development and Education, University of Amsterdam, the Netherlands

ARTICLE INFO

Keywords:
Achievement
Higher education
Intervention
Meta-analysis
Self-regulated learning

ABSTRACT

It is often assumed that interventions aimed at supporting students' self-regulated learning (SRL) are effective for improving achievement because these interventions support SRL activity. In this study, meta-analytic structural equation modeling (MASEM) was used to test whether SRL activity indeed mediates the effect of SRL interventions on achievement in higher education. Contrary to popular belief, the results only provide evidence for partial mediation. Furthermore, three separate meta-analyses were performed to investigate the role of possible moderators of the relations between: (1) SRL interventions and achievement, (2) SRL interventions and SRL activity, and (3) SRL activity and achievement. Although SRL interventions were effective in improving SRL activity and achievement, most of the study, measurement, and intervention moderators did not explain significant variance of the investigated effect sizes. Other factors, such as task motivation and time on task, potentially influence the effectiveness of SRL interventions. Practical, theoretical and methodological implications are provided.

1. Introduction

Factors influencing academic achievement have always received great attention from scholars, and self-regulated learning (SRL) is no exception (Boer, Donker-Bergstra, Kostons, & Korpershoek, 2013). Higher education students who engage in SRL are actively involved in their learning process. As SRL activity has consistently been found to be related to student achievement, research into ways to support SRL is plentiful (e.g., Azevedo & Cromley, 2004; Azevedo, Cromley, & Seibert, 2004; Bannert, Hildebrand, & Mengelkamp, 2009; Broadbent & Poon, 2015; Nietfeld, Cao, & Osborne, 2006; Stark & Krause, 2009). In general, these SRL interventions that support students' knowledge of SRL and their engagement in SRL activities, show positive results on achievement (e.g., Boer et al., 2013; de Bruijn-Smolders, Timmers, Gawke, Schoonman, & Born, 2016; Dignath & Büttner, 2008). It is assumed that SRL interventions are effective in improving achievement due to their effects on students' engagement in SRL activities: the SRL intervention improves students' use of SRL activities and this improvement leads to better performance. Proof that engagement in SRL activities mediates the effect of SRL interventions on achievement is nevertheless lacking. The first goal of this paper is therefore to perform a mediation analysis with meta-analytic data, to test if engagement in SRL activities is a significant mediator of the effect of SRL interventions on achievement. Furthermore, it is not yet known what makes an SRL intervention effective for improving students'

* Corresponding author. Department of Education, Utrecht University, P.O. Box 80140, 3508, TC, Utrecht, Netherlands.
E-mail address: R.S.Jansen@uu.nl (R.S. Jansen).

https://doi.org/10.1016/j.edurev.2019.100292
Received 3 September 2018; Received in revised form 17 June 2019; Accepted 10 September 2019
Available online 16 September 2019

SRL activities or achievement. The second goal of this paper is therefore to test which characteristics influence the effectiveness of SRL interventions. To do so, separate meta-analyses are conducted of (1) the effect of SRL interventions on achievement, (2) the effect of SRL interventions on SRL, and (3) the relation between engagement in SRL activities and achievement.

1.1. Self-regulated learning

SRL is central to students' learning process, as students that engage in SRL take control of their own learning process. Multiple models and frameworks of SRL exist (Pintrich, 2000, 2004; Winne & Hadwin, 1998; Zimmerman, 1990, 2002), but the main components of the models are similar (Pintrich, 2000; Puustinen & Pulkkinen, 2001). Students who self-regulate their own learning process are metacognitively, behaviorally, and motivationally active in their learning and they proceed through three phases: a preparatory phase, a performance phase, and an appraisal phase (Zimmerman, 1986, 2002). In the preparatory phase, students prepare for the learning task at hand, they plan their work, and set goals. In the performance phase, students engage in cognitive strategies to learn the material at hand, they monitor their learning, regulate their learning strategies, and they allocate their resources (e.g., time and help) in the most efficient manner. Lastly, in the appraisal phase, students reflect on their learning and determine which strategies were effective and what they could do differently the next time they study (Pintrich, 2000; Puustinen & Pulkkinen, 2001; Zimmerman, 2002).

In each phase, learners may engage in different activities to regulate their learning. Students' engagement in these different activities is captured under the umbrella term engagement in SRL activities. Students that are metacognitively involved in their learning strategically plan, monitor, and reflect on the cognitive strategies they use for learning. The SRL activities they engage in are goal setting and strategic planning before learning (preparatory phase), comprehension monitoring and strategy regulation during learning (performance phase), and reflection after learning (appraisal phase; Puustinen & Pulkkinen, 2001; Zimmerman, 2002). Next to these metacognitive activities, SRL also entails behavioral regulation and motivational regulation (Zimmerman, 1986). Behavioral regulation includes the SRL activities time management, environmental structuring (i.e., studying in a suitable location), and help seeking. Motivation regulation includes persisting when motivation drops, which has also been termed effort regulation (Pintrich, 1999). Motivational regulation and behavioral regulation both mostly take place during learning (performance phase) and can together be classified as resource management activities, as these activities all focus on keeping resources (e.g, help, environment, attention) at an appropriate level (Pintrich, 1999).

SRL models differ in their explanation of the role of motivation (Panadero, 2017). Some consider motivation for learning part of SRL (e.g., Boekaerts, 1992; Zimmerman, 2002, 2008), while others consider persistence and regulating motivation part of SRL, but consider task motivation a precursor for successful SRL (e.g., Efklides, 2011; Ning & Downing, 2012; Pintrich, 1999, 2000; Schunk, 2005). In the latter models, it is reasoned that learners will not engage in successful SRL without sufficient motivation for the task at hand. We adhere to this latter definition, leaving motivation for learning (i.e., task value, goal orientation, and self-efficacy) outside of the scope of the current review.

1.2. The importance of SRL for student achievement

Students who self-regulate choose the most effective cognitive learning activity (e.g., rehearsal, elaboration, note-taking) depending on the current learning task and broader context (Boekaerts, 1992; Winne & Hadwin, 1998; Zimmerman, 2002). Students' SRL activities thereby influence their cognitive activities (Nelson & Narens, 1990), and students who actively self-regulate their learning thus engage in more effective cognitive strategies. Students' progress during learning is a result of their use of cognitive strategies (Dunlosky, Rawson, Marsh, Nathan, & Willingham, 2013). This progress is, in turn, used as input for further self-regulation through monitoring (Beishuizen & Steffens, 2011; Nelson & Narens, 1990). SRL and the use of cognitive strategies thus form a cyclical process during learning (Nelson & Narens, 1990; Veenman, 2016). The influence of SRL on the use of cognitive learning strategies likely explains why prior review studies consistently show that SRL is related to higher student achievement (Boer et al., 2013; de Bruijn-Smolders et al., 2016; Dignath, Buettner, & Langfeldt, 2008; Dignath & Büttner, 2008; Sitzmann & Ely, 2011). Cognitive strategies thus influence the effects of SRL on achievement. In the current review we however focus on the effects of SRL interventions, thereby placing engagement in cognitive learning activities outside of our scope.

Due to the importance of SRL for student achievement, numerous studies have been conducted on how academic achievement is affected by interventions that aim at supporting students' engagement in SRL activities (e.g., Azevedo & Cromley, 2004; Azevedo et al., 2004; Bannert et al., 2009; Broadbent & Poon, 2015; Nietfeld et al., 2006; Stark & Krause, 2009). These so-called *SRL interventions* take many forms. They are aimed at supporting students' engagement in SRL activities, either by supporting the quality of students' engagement in SRL activities, the quantity of students' engagement in SRL activities, or both. Examples of SRL interventions include informing students' about effective SRL activities and their importance, or prompting students at the end of lecture activities to reflect on the course material, on the completed assignments, and on the strategies they used for learning.

1.3. The mediating role of SRL activities

Empirical studies exploring the effect of SRL interventions are usually focused on improvements in achievement. They are often based on the assumption that SRL interventions stimulate students' engagement in SRL activities (relation 2 in Fig. 1), in turn leading to increased academic achievement (relation 3 in Fig. 1). An indirect effect of the interventions on achievement is thus assumed. However, to the best of our knowledge, no studies have tested the mediating role of SRL activities, and only very few studies included

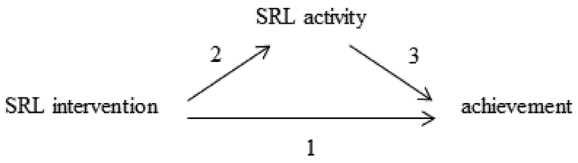

Fig. 1. The effect of SRL interventions on achievement, mediated by SRL activities.

all three relations between SRL interventions, SRL activities, and achievement (Garavalia & Gredler, 2002; Greene, Hutchison, Costa, & Crompton, 2012; Schmidt & Ford, 2003) in their analyses. While some studies included both the effect of SRL interventions on SRL activities as well as the effect of SRL interventions on achievement, none of these studies tested the mediating role of SRL activities on the effect of SRL intervention on achievement. The majority of studies examined the direct relationship between SRL interventions and academic achievement (relation 1 in Fig. 1; e.g., Azevedo et al., 2004; Bannert, Sonnenberg, Mengelkamp, & Pieger, 2015; Lusk, 2016).

Similarly, existing meta-analyses and systematic reviews focused only on parts of the framework presented in Fig. 1, namely: the effect of SRL interventions on achievement (Boer et al., 2013; de Bruijn-Smolders et al., 2016; Devolder, van Braak, & Tondeur, 2012; Dignath & Büttner, 2008; Dignath et al., 2008; Hattie, Biggs, & Purdie, 1996), the effect of SRL interventions on SRL activity (de Bruijn-Smolders et al., 2016; Dignath & Büttner, 2008; Dignath et al., 2008; Hattie et al., 1996), and the relationship between SRL activity and achievement (Broadbent & Poon, 2015; Dent & Koenka, 2016; Sitzmann & Ely, 2011). A meta-analysis that focuses on the mediating role of SRL in the effect of SRL interventions on achievement is missing. The first goal of this paper is therefore to systematically review SRL as a mediator of the effect of SRL interventions on achievement, in the context of higher education. We will perform a meta-analytic mediation analysis to test the commonly made claim that SRL interventions are effective for improving student achievement by improving students' engagement in SRL activities.

Additionally, most existing reviews were conducted in primary and secondary education. Only a few reviews include data from students in higher education: Sitzmann and Ely 2011; higher education and employees), Broadbent and Poon 2015; online higher education), De Bruijn et al. (2016; higher education), and (Hattie et al. 1996; primary, secondary, and higher education and employees). This lack of overview studies aimed at higher education is unfortunate because SRL is of great importance for student achievement in higher education. In higher education, learners are provided greater autonomy as there is limited external regulation of their learning process (Beishuizen & Steffens, 2011; Sitzmann & Ely, 2011). Learners for instance study mostly outside of the classroom at their own pace and attendance is often not required. The increased independence of learners leads to a greater need for learners to take control of their own learning process in higher education compared to primary and secondary education making SRL more important (Beishuizen & Steffens, 2011; Wang, Shannon, & Ross, 2013). The lack of reviews focusing on SRL in higher education, combined with the importance of SRL in this context, leads us to focus on SRL in higher education in the current review.

1.4. What makes SRL interventions effective?

Due to the importance of SRL for student achievement, numerous interventions have been developed to support students' SRL activity. However, there are currently few guidelines available on how to develop effective SRL interventions in higher education. The second goal of this paper is therefore to systematically review what characteristics of SRL interventions lead to improved achievement (relation 1 in Fig. 1) and SRL activity (relation 2). Thus, while we group all intervention studies in the first part of the paper, ignoring differences between studies to calculate the mediation by SRL activity, in the second part of the paper we focus on the differences between studies by performing moderator analyses. Moderator analyses of the effect of SRL interventions could provide clues on how to design effective SRL interventions to support achievement and/or SRL activity. Such guidelines would be valuable for educational practice. In the second part of the paper, we therefore perform meta-analyses of the three relations outlined in Fig. 1. The relation between SRL activity and achievement (relation 3) is included in the current study to enable calculation of the mediation of the effect of SRL interventions on achievement by SRL activity; without inclusion of the relation between SRL activity and achievement this would not be possible. However, there already is an abundance of research consistently showing this relation to be significant and positive (Boer et al., 2013; de Bruijn-Smolders et al., 2016; Dignath & Büttner, 2008; Dignath et al., 2008; Sitzmann & Ely, 2011).

In the next section, we discuss existing knowledge about the effect of SRL interventions on engagement in SRL activities and on achievement. As there is ample research into SRL, SRL interventions, and achievement, we provide an overview of the current literature based on existing review studies. Even though the reviews discuss the three relations in isolation, and most reviews are conducted in primary and secondary education, they offer hypotheses of what might constitute an effective SRL intervention in higher education to improve both students' engagement in SRL activities and their achievement. Analyzing the existing knowledge on SRL interventions thereby provides indications for characteristics to include as moderators in the current study.

1.5. Potential moderators of the effectiveness of SRL interventions

Not all SRL interventions have been found to be equally effective in supporting students' engagement in SRL activities and students' achievement (Boer et al., 2013; Dignath & Büttner, 2008; Dignath et al., 2008). Characteristics that potentially influence the

measured effectiveness of SRL interventions on SRL activity and achievement have been identified by analyzing the differences between empirical SRL intervention studies. Characteristics included in previous SRL review studies have furthermore been incorporated in the current review. These characteristics are used as moderators in the present study and are described below. We distinguish between five intervention characteristics and three measurement characteristics.

The first characteristic for which interventions show variety is whether *cognitive strategies are included,* so whether the intervention focuses solely on SRL, or also on cognitive strategies. Students must possess an adequate repertoire of cognitive strategies in order to be able to successfully self-regulate; without such a repertoire students cannot regulate their learning behavior as they have insufficient strategies to choose from (Hattie et al., 1996; Paris & Paris, 2001). It is likely that the need for instruction about cognitive strategies decreases when students reach higher levels of education, because by then, they have internalized more cognitive strategies. Students in higher education, therefore, may benefit more from interventions that aim at supporting their SRL activities, compared to strategies aimed at supporting their cognitive strategies. For instance, they would benefit more from support on how to plan their work (an SRL activity) than how to memorize facts (a cognitive strategy). As some studies include cognitive strategies in their SRL intervention (e.g., Dörrenbächer & Perels, 2016; Wischgoll, 2016), while others do not (e.g., Greene et al., 2012; Nietfeld et al., 2006), the inclusion of cognitive strategies is incorporated as a moderator in the meta-analysis to test whether it influences the effectiveness of SRL interventions.

The second intervention characteristic is the *format of the intervention.* Students need to have knowledge about SRL and SRL activities in order to be able to successfully engage in SRL (Dignath & Büttner, 2008; Paris & Paris, 2001; Zimmerman, 2002). In primary education, Dignath and Büttner (2008) found that SRL interventions were most effective for improving students' achievement when they contained instruction on SRL as a concept and on possible SRL activities students could engage in. In secondary education, interventions with metacognitive *instruction* were less effective for improving students' achievement compared to interventions with metacognitive *reflection* (Dignath & Büttner, 2008). Thus, instead of instruction on metacognition and metacognitive activities, students benefitted most from interventions that stimulated them to reflect on their learning process and to engage in the metacognitive activities they already knew. Students in secondary education possess more knowledge on SRL activities than students in primary education and can, therefore, benefit from metacognitive reflection (Dignath & Büttner, 2008). For instance, while students in primary education benefit from instruction on what a planning is and how one should construct a planning, students in secondary education already know what a planning is but need to be stimulated to create a planning and monitor their progress. Students in higher education may have internalized even more knowledge on SRL and SRL activities. This would imply that SRL interventions for successfully improving students' achievement in higher education should focus on stimulating students to reflect on their cognitive strategy use and activate students' existing knowledge of SRL (Bannert & Reimann, 2012). This could for instance take the form of scaffolding, where students are for example prompted to self-explain their steps to mastery of the learning material until they no longer need to be prompted (Van Laer & Elen, 2017). Thus, whether the intervention provides instruction, application, or prompting of SRL will be included as a moderator to test potential differences in the effectiveness of interventions based on their format.

The third characteristic for which interventions show variety is the *timing of the intervention* in relation to the learning context. In some intervention studies, the intervention is placed before learning (e.g., Azevedo & Cromley, 2004). In other studies, the intervention is incorporated during the whole lab experiment or course (e.g., Delen, Liew, & Willson, 2014), or only in the second half (e.g., Nückles, Hübner, & Renkl, 2009). Incorporating the SRL support before the learning task allows learners to benefit from the support during the whole task, which may lead to larger effects. On the other hand, incorporating the intervention at a later stage allows students to first get accustomed to the learning environment, which may reduce the possibility of overwhelming them with too much information. As there are thus differences in the timing of the intervention, and these differences may influence the effectiveness of the SRL interventions, timing is included as a moderator.

The fourth intervention characteristic that will be included as a moderator is whether the intervention is *tailored to the learning context.* As the effectiveness of SRL activities is partly context dependent, interventions should aim at helping students determine what the effective SRL activities are in the specific context at hand (Puustinen & Pulkkinen, 2001). Students can benefit from the synergy between domain-learning, cognitive strategy instruction, and SRL activities (Perels, Gürtler, & Schmitz, 2005). For this reason, domain-general prompts for SRL, asking students to reflect on their learning, have been found to be less effective for supporting students' SRL behavior than domain-specific prompts in which students, for instance, were asked how well they comprehended a specific concept within the domain (Devolder et al., 2012). To test whether this also holds in higher education, whether the intervention is tailored to the learning context (i.e., domain-specific) will be included as a moderator.

The fifth intervention characteristic is *the type of SRL activity* supported by the intervention. Broadbent and Poon (2015) and Sitzmann and Ely (2011) found diversity in the magnitude of the correlation between students' reported engagement in a range of SRL activities and student achievement in higher education. For instance, effort regulation in the form of persistence was found to be much more strongly related to achievement than help-seeking (Sitzmann & Ely, 2011). Some SRL activities are thus more strongly associated with achievement than others. This difference in the strength of the correlation between SRL activities and achievement may also explain why interventions aimed at some SRL activities are more beneficial for academic achievement than others (Boer et al., 2013). The question remains whether SRL interventions should focus on a subset of particular SRL activities (that are strongly related to achievement), or that interventions should focus on supporting students' SRL in general; and if the preferred focus is dependent on the goal of the SRL intervention (improving achievement, relation 1 in Fig. 1, versus supporting SRL activity, relation 2). Therefore, the SRL activity the intervention is aimed at will be included as a moderator to explore whether SRL interventions that focus on different aspects of SRL have different effects. All metacognitive (e.g., planning, comprehension monitoring) and resource management (e.g., time management) activities introduced will be included. More information about the SRL activities included and

how they are coded will be presented in the Method section.

The *type of SRL activity measured* is the first measurement characteristic that will be included as a moderator. It is included to see if the results of Broadbent and Poon (2015) and Sitzmann and Ely (2011), showing that some SRL activities are more strongly related to achievement than others, are replicated in the current meta-analyses.

The second measurement characteristic included as a moderator is the *instrument* used to measure SRL. The validity of instruments used to measure SRL is widely debated (e.g., Greene & Azevedo, 2010; Panadero, Klug, & Järvelä, 2016; Veenman, 2016; Veenman, Van Hout-Wolters, & Afflerbach, 2006; Winne, 2010). By incorporating *the instrument used to measure students' engagement in SRL activities* we assess whether differences in the effects of SRL interventions on SRL activity, and differences in the relationship between SRL activity and achievement, are due to the instrument used to measure SRL.

Finally, the employed *achievement measure* is the third measurement characteristic included as a moderator. Different achievement measures are in use to test the effects of SRL interventions and the relation between SRL activity and achievement, including course grade (e.g., Cleary, Callan, Malatesta, & Adams, 2015; van den Boom, Paas, & van Merriënboer, 2007) and GPA (e.g., Masui & De Corte, 2005; McKenzie, Gow, & Schweitzer, 2004). By including the employed achievement measure as a moderator, we are able to test whether the achievement measure can explain variability between studies.

1.6. Present study

In the present study, we systematically review empirical research on the effectiveness of SRL interventions on students' engagement in SRL activities and on students' academic achievement. We first perform a meta-analysis on SRL activities as mediating the effect of SRL interventions on achievement. The research question we pose is: *To what extent is the effect of SRL interventions on students' achievement due to students' SRL activity?* By calculating the mediating effect of SRL in the relationship between SRL interventions and academic achievement (relations 2 and 3 in Fig. 1), we determine whether SRL interventions indeed lead to increased achievement through changes in students' SRL activity. We hereby aim to supply empirical evidence for this assumption.

Second, existing reviews (Boer et al., 2013; Dignath et al., 2008; Dignath & Büttner, 2008) have provided indicators of what constitutes an effective SRL intervention for supporting students' SRL activity and students' academic achievement in primary and secondary education. Unfortunately however, in line with the empirical research in the field of SRL, the reviews include only part of the framework presented in Fig. 1. A comprehensive meta-analysis including all three relations at once is still missing. We will therefore perform meta-analyses of (1) the effectiveness of SRL interventions for improving student achievement, (2) the effectiveness of SRL interventions for supporting students' SRL activity, and (3) the relationship between SRL activity and achievement (Fig. 1). The meta-analyses will be conducted for the context of higher education, as SRL becomes more important with an increase in autonomy, as is common in higher education (Beishuizen & Steffens, 2011). The research questions we thus pose are: *Which factors significantly influence the effect of SRL interventions on achievement? Which factors significantly influence the effect of SRL interventions on SRL activity?* and *Which factors significantly influence the relationship between SRL activity and achievement?*

2. Method

To answer the research questions, two types of analyses were conducted. First, a mediation analysis was conducted using *meta-analytic structural equation modeling* (MASEM; Jak, 2015) to test the extent to which the effect of SRL interventions on achievement is due to changes in students' engagement in SRL activities. Second, the overall effect of SRL interventions on academic achievement (relation 1), the overall effect of SRL interventions on SRL activity (relation 2) and the relation between SRL activity and achievement (relation 3) were calculated, after which the variation within each of these relations was further explored. For each relation separately, factors that significantly explain variance in effect sizes were identified through moderator analyses. Thus, in addition to the mediation analysis using MASEM, three separate meta-analyses were conducted, each with complementing moderator analyses.

2.1. Literature search

To identify studies to include in the meta-analyses, a literature search was conducted on July 8th, 2016. The search queries included key words to narrow down the search to educational contexts and to exclude studies conducted in primary or secondary education. The search queries were kept broad on purpose to avoid missing relevant literature. This mainly concerned terms relating to SRL, as there are numerous terms in use in the educational literature describing SRL or SRL activities. There were no restrictions concerning date of publication; all hits registered in the databases until the search date were considered. Search terms for the relation between SRL intervention and SRL activity (relation 2 in Fig. 1) were *education/learn*/class* AND scholar*/student*/*grad*/pupil*/employee*/learner*/participant* AND SRL/SDL/self-reg*/self-dir*/self reg*/self dir*/metacogniti*/study skill*/study strateg* AND intervention*/experiment*/condition*/treatment*/compar*/train*/support*/improve*/*

scaffold/effect* of AND NOT primary education/kindergarten/young child*/young-child* AND NOT gifted/disab*/disorder.* Searches were conducted in the PsycInfo, Scopus and Web of Science databases searching in title, abstract, and keywords. This search resulted in 4,236, 4,157, and 3,512 hits respectively. For the relation between SRL activity and achievement (relation 3 in Fig. 1) the search terms were *education/learn*/class* AND scholar*/student*/*grad*/pupil*/employee*/learner*/participant* AND SRL/SDL/self-reg*/self-dir*/self reg*/self dir*/metacogniti*/study skill*/study strateg* AND learning outcome*/achievement/*perform*/success*/grade AND NOT primary education/kindergarten/young child*/young-child* AND NOT gifted/disab*/disorder.* This search resulted in 5,041 (PsycInfo), 5,009 (Scopus), and 4,290 (Web of Science) hits. For the relation between SRL interventions and achievement (relation 1 in Fig. 1) no

Table 1
Overview of inclusion and exclusion criteria.

Category	Inclusion	Exclusion
Language	Article written in English	Article not written in English
Population	Representative sample	At risk students. Low/high performing students. Students with a disorder/disability
Setting	Self-regulation in an educational context (i.e. self-regulated learning)	Self-regulation in a different context (e.g., self-regulation of alcohol consumption)
Intervention	Control group (simultaneous or non-simultaneous)	No control group
	The intervention must be aimed at applying self-regulation in an educational context	Interventions not directly aimed at improving SRL (e.g., meditation)
	Description of the intervention in the study	Description is absent, making it impossible to infer the content of the intervention
SRL measurement	Measure must solely incorporate (aspects of) SRL	Measures that combine SRL measurement with other measures (e.g., cognitive measures)
	Positive SRL behaviors	Negative SRL behaviors (e.g., avoidance of help seeking)
Achievement	Achievement measured with an individual task	Achievement measured with a collaborative task
Intervention studies	The intervention takes place before the achievement measure	The intervention takes place after the achievement measure
Correlational studies	The SRL measure precedes the achievement measure, or the SRL measure and the achievement measure relate to the same educational context	The achievement measure precedes the SRL measure and does not relate to the same educational context
Statistics	Quantitative studies	Qualitative studies

separate literature search was conducted as this would have resulted in a subset of the articles already retrieved by the other two literature searches. For all three relations combined, there was a total of 26,245 hits. After removal of duplicates, 13,259 studies remained.

Titles and abstracts of the found studies were filtered based on a predefined list of inclusion and exclusion criteria, which can be found in Table 1. Some of these criteria were already included in the search terms to reduce the number of hits, for instance by excluding articles containing the words *gifted, disab** and *disorder* in the title, abstract or keywords. A total of 700 studies was retained after filtering based on titles and abstracts. Only the first identified reason for exclusion was recorded. The most often identified reason for exclusion was that the study did not focus on SRL as defined in the Introduction or that it did not fit any of the three relations. Exemplary studies that were excluded for this reason focused on the effect of logical reasoning exercises on self-control (Bertrams & Schmeichel, 2014) or the validation of SRL-related questionnaires (Credé & Phillips, 2011). Other common reasons for exclusion were (a) other context than learning/education, (b) participants at school level below higher education (mostly studies from high school contexts), and (c) qualitative studies. For all remaining studies (journal articles, conference proceedings, and dissertations), we attempted to obtain a full text version. Full text versions were further screened based on the inclusion and exclusion criteria. Special attention was paid to the availability of statistical information necessary to calculate effect sizes.

The final set of included studies contained 126 articles. The most common reasons for articles to be excluded were (a) that the article could not be obtained, mostly because it was a conference abstract for which no full version existed, (b) that there was no measure of SRL or achievement, and (c) that the intervention did not focus on improving SRL. For relation 1 (SRL intervention – achievement), 51 articles were included; for relation 2 (SRL intervention – SRL activities), 32 articles were included; and for relation 3 (SRL activities – achievement), 70 articles were included. Articles could include data on one or multiple relations. Articles that included data on multiple relations were included in all meta-analyses for which they provided sufficient information. Thus, a study could potentially be included in all three meta-analyses if it provided information about the effect of an SRL intervention on achievement, about the effect of an SRL intervention on students' engagement in SRL activities, and about the relationship between students' SRL activity and their achievement. Each study could furthermore include multiple study samples and multiple effect sizes. For example, a study may measure multiple SRL activities, and correlate each of them separately to students' grades. If the study also included measures that we do not consider part of SRL (e.g., self-efficacy), effect sizes on these measures were excluded from the meta-analyses. If the SRL measure reported was a combination of aspects that we do and do not consider part of SRL, the effect size was excluded as it was not a measure that captured solely SRL. A reference list with all studies included in one or multiple of the meta-analyses is available as supplementary material (online only). A visual overview of the literature selection process is presented in Fig. 2.

2.2. Coding of included studies

All included studies were coded. When possible, the coded characteristics were based on existing knowledge of the effect of SRL interventions. The coding framework consisted of five segments and was developed in such a way that it was suitable for coding the different types of studies. The first and second segment were relevant for all studies and were labeled *educational context* and *sample characteristics*. For the educational context, we coded the type of study (lab or real course), the setting in which the study was conducted (online or offline), and the academic subject. For the sample characteristics, we coded participants' educational level, age, gender, and region.

The third segment focused on the *independent variable* of the study. For relations 1 (SRL intervention → achievement) and 2 (SRL

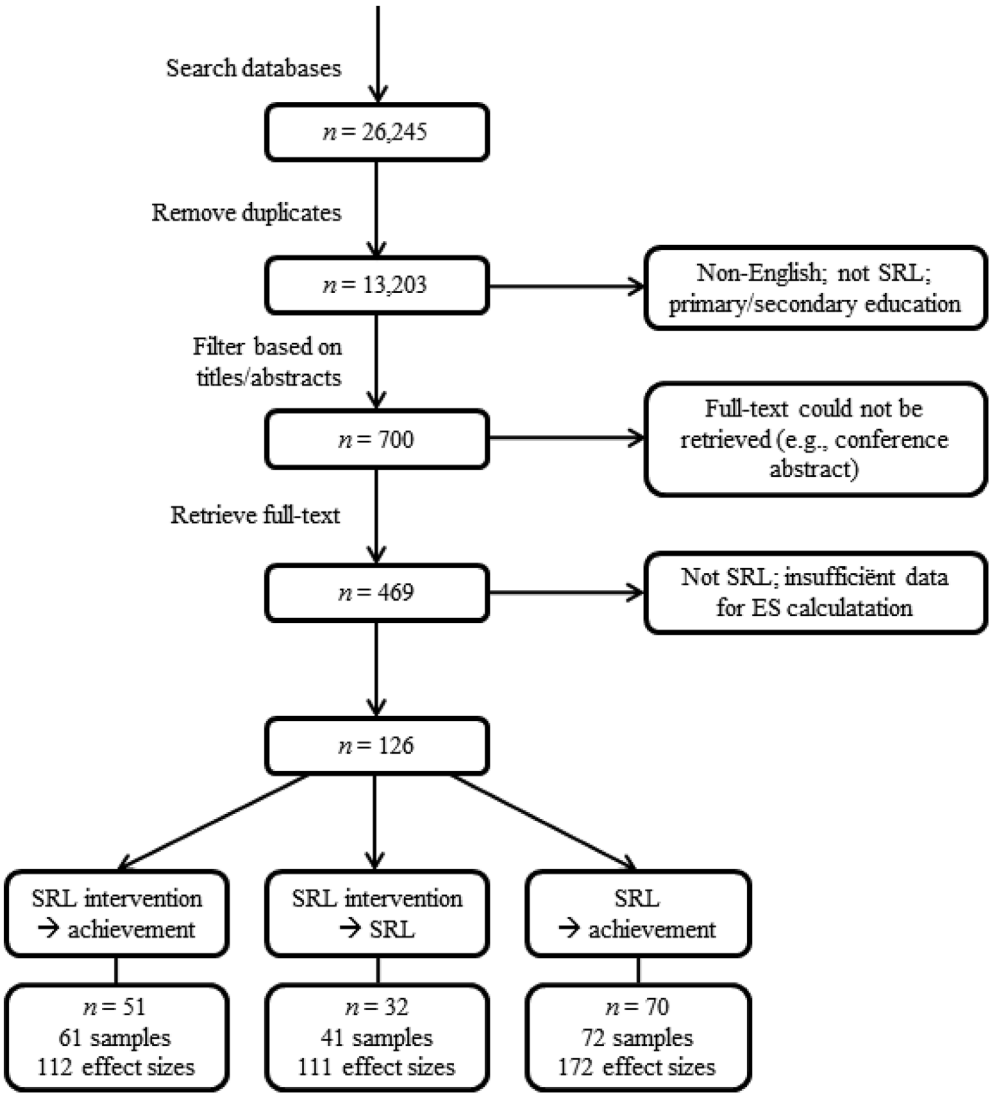

Fig. 2. Overview of the literature selection process.

intervention → SRL) the independent variable was the SRL intervention. For relation 3 (SRL → achievement) it was SRL. For SRL interventions, we coded the length of the educational context, the length of the intervention (in weeks), the duration of the intervention (e.g., actual training hours), the timing of the intervention relative to the learning task, the inclusion of cognitive strategy training, the format, the timing of the intervention relative to the learning task, whether the intervention was tailored to the task, and the type of SRL activity supported. We furthermore coded the type of control group, the manner in which students were allocated to groups, and group equality to determine study quality. For SRL measurements, we coded the instrument used to measure SRL, the reliability of the measure, and the type of SRL measured.

The fourth segment of the coding framework focused on the *dependent variable*. For relations 1 and 3 the dependent variable was achievement, while for relation 2 it concerned SRL. The coding of SRL as a dependent variable for relation 2 was identical to the coding of SRL as an independent variable for relation 3. For achievement we coded the achievement measure used. Finally, the fifth segment was labeled *timing of measures*, including the time between the SRL intervention and the measurement of achievement. Thereby all intervention and measurement characteristics presented in the Introduction (section 1.5) were included in the coding scheme. The complete coding framework is available as supplementary material (online only).

Reliability of the coding framework was assessed in two rounds. In the first round, two raters independently coded a random sample of 20 studies (10 studies by the first and second author, and 10 by the first and third author). For most of the variables in the coding scheme, not all options were equally likely, which means Cohen's kappa was not suitable as a measure of reliability (Feinstein & Cicchetti, 1990). Instead, percentage agreement was used as the reliability measure. The percentage agreement ranged between 50% and 100% for the entire coding scheme. Variables that were coded with insufficient reliability were discussed among the first three authors. As a result of this discussion, the coding options for two variables were adjusted to better suit the characteristics of the

Table 2
Included moderators.

Category	Coded information
Study characteristics	Academic subject
	Educational setting (online or offline/web-based)
	Study design quality
	Context
Measurement characteristics	SRL activity measured
	Measurement instrument SRL
	Achievement measure
Intervention characteristics	Inclusion of cognitive strategies
	Format
	Timing
	Tailored to the learning context
	SRL activity supported

interventions as found in the sample set. A second round of coding was conducted with a new random sample of 21 studies (11 studies were coded by the first and second author, and 10 by the first and third author). The studies in this second sample set were all intervention studies (relation 1 and 2 in Fig. 1). A sufficient agreement percentage ranging between 80% and 100% was found for the variables that addressed studies on the relationship between SRL and achievement (relation 3 in Fig. 1) in round 1. It was therefore not necessary to include any more studies on this relationship in round 2. Interrater agreement for this second sample set ranged between 81% and 100%. The only exception was the type of SRL activity supported by the intervention with an interrater agreement of 71%. The discrepancies in coding had two main causes. First, in some primary studies, the descriptions of the intervention were unclear. Second, in some of the primary studies, the terminology used to label the intervention was different from the labeling with our coding scheme when considering the content of the intervention. These issues complicated accurate labeling of the type of SRL activity supported by the intervention. As this variable is crucial for the current meta-analyses, we decided that this variable would be coded by two authors for all studies. Disagreements on the coding of this variable were resolved by discussion. The interrater agreement for each variable is available in the coding framework in the supplementary materials (online only).

2.3. Moderators

Based on the coding scheme, moderators were included in the conducted meta-analyses on the effect of SRL interventions on achievement (relation 1), the effect of SRL interventions on SRL activity (relation 2), and the relationship between SRL activity and achievement (relation 3). Three groups of moderators were included: study characteristics, measurement characteristics, and intervention characteristics. An overview of the included moderators can be found in Table 2.

2.3.1. Study characteristics

Academic subject of the study material was included as a moderator and was classified as social sciences, humanities, formal sciences, applied sciences, or a combination of these. The *educational setting* in which the study was conducted, either online or offline/web-based was included as the second moderator. The third moderator was the *quality* of the study. The quality of the studies was determined based on three coded aspects: whether the experimental and control group were simultaneous or non-simultaneous (i.e., same or different cohorts), how students were allocated to groups, and whether groups were equal on relevant covariates. Three quality levels were created. Studies of the highest quality had a simultaneous control and intervention group, students were randomly assigned to conditions, and groups were equal. Studies of the second level of quality had a simultaneous control group, but students were either allocated based on pre-existing groups or not-random at all. In all studies at level two, it was either assured by the authors of the primary study that groups were equal or inequality was statistically controlled for by the authors of the primary study. The two studies with a non-simultaneous, non-random control group were also classified as quality level 2, as the authors of these studies assured that the control and intervention groups were equal. The third and lowest level of quality contained studies for which information about group equality was missing from the article. The fourth and final study characteristic included as a moderator was study *context*, which was either lab or real course. The educational grade level of students was not included as a moderator as there was too little diversity in this variable across studies.

2.3.2. Measurement characteristics

For studies in which SRL was measured (relation 2 and 3 in Fig. 1) the *instrument* used to measure SRL and *the SRL activity or set of SRL activities measured* were included as moderators. SRL measurements were coded as either a questionnaire, a counted measure (a combination of think aloud and counting specific, observed behaviors), or self-assessment accuracy (the difference between a students' self-assessed score and the obtained score). The SRL activity that was measured was coded as one of fourteen values: measures focusing on a specific metacognitive aspect (6 values: planning, strategy regulation, goal setting, monitoring, reflection, and self-assessment), measures focusing on a specific resource management aspect (5 values: help seeking, environment structuring, time management, environment structuring and time management, and persistence), measures focusing on metacognition in general (1 value), measures focusing on resource management in general (1 value), and measures focusing on both metacognition and resource

management (1 value). The type of SRL measured was coded based on our knowledge of the questionnaire included, example items included in the primary study, or the description of the measure in the primary study. As some values were coded only rarely, and to decrease the number of moderator dummies in the analyses, the measures were clustered into larger groups of SRL measures. These groups were: single aspect metacognitive measures, multiple aspect metacognitive measure, single aspect resource management measures, multiple aspect resource management measures and finally general measures incorporating both aspects of metacognition and resource management. For studies with achievement as the dependent variable, the employed achievement measure was included as a moderator. Achievement measures were classified as task performance, course performance, GPA, or performance on a transfer task.

2.3.3. Intervention characteristics

Five intervention characteristics were included as moderators for the meta-analysis of the effects of SRL interventions on achievement and the meta-analysis of the effects of SRL interventions on SRL. The *SRL activity* or *the set of SRL activities* supported by the intervention was included as a moderator. The SRL activities were clustered in the same manner as the measurement characteristic *SRL activity measured*. As there were no interventions that focused on a single aspect of resource management, or on resource management in general, three categories of SRL support remained: interventions aimed at a single metacognitive activity, interventions aimed at multiple metacognitive activities, and general SRL interventions aiming at both metacognitive and resource management activities. The second moderator was *the inclusion of cognitive strategies*, which indicated whether the intervention included cognitive strategy instruction. The third moderator indicated the *timing* of the intervention in relation to the learning context: before the learning phase, during the first half of learning, during the second half of learning, during the whole learning task, or both before and during learning. The fourth moderator was whether or not the intervention was *tailored* to the learning context. In tailored interventions, the SRL instruction or the SRL activity students had to engage in was coupled to the learning task, instead of a stand-alone instruction independent of the learning task. The final included moderator was the *instructional format* of the intervention. This moderator was recoded into three variables, each indicating whether a specific aspect was present in the intervention: instruction (yes/no), application (yes/no), and prompts (yes/no). It was planned to also include intervention intensity as a moderator by combining information on the length and duration of the intervention. This information was however reported in only very few articles and could therefore not be used as an indicator of study intensity.

2.4. Effect size extraction

After coding the studies, effect sizes were extracted. For the intervention studies (relation 1 and 2), Cohen's *d* was used as a measure of effect size. The preferred method to calculate Cohen's *d* was to use the means and standard deviations of posttest scores reported in the study, but if only other statistics, like means and standard deviations of the gain scores (difference between pre- and posttest) were reported, those were used instead.[1]

The studies included for relation 3 (SRL activities → achievement) were all correlational in nature. Pearson's *r* was used as a measure of effect size. In most cases, the correlation coefficient between a measure of SRL and achievement was reported and could be incorporated in the current meta-analysis. In a few cases, only the regression analyses or path coefficients from a structural equation model were reported. In these situations, the standardized betas or path coefficients were included.[2]

After calculating all effect sizes, special attention was paid to the direction of the effect sizes concerning self-assessment accuracy (one of the measures of SRL). If the accuracy of self-assessment increases, the distance between the predicted score and the true score decreases. If students with higher accuracy obtain higher achievement scores, the effect size is negative. In these cases, the sign of the effect size was reversed. The relation between self-assessment and achievement was hereby coded in the same direction as all other relations: positive effect sizes indicated that a higher SRL value related to higher achievement.

2.5. Analyses

First, descriptive statistics concerning the number of extracted effect sizes were calculated. Next, the three sets of effect sizes (one for each relation) were separately checked for indicators of potential publication bias. Then, the potential mediating effect of SRL activity on the relationship between SRL interventions and achievement was evaluated with meta-analytic structural equation modelling (MASEM). With MASEM it is possible to combine the available information from separate studies and to test the mediation model presented in Fig. 1, even though only three studies included statistical information on all three relations. MASEM combines the benefits of structural equation modeling (SEM) and meta-analysis (Jak, 2015). For the MASEM analysis, the random-effects Two Stage SEM approach was used (TSSEM; Cheung, 2014). TSSEM employs multivariate random-effects meta-analysis to pool correlation matrices at Stage 1. Then, at Stage 2, weighted least squares estimation with the inversed asymptotic covariance matrix of the Stage 1

[1] For intervention → achievement 12 of the 112 effect sizes could not be calculated based on the means and standard deviations of posttest scores. The way the effect size was calculated (posttest scores or another method) was found not to be a significant moderator ($p = .466$). For intervention → SRL 10 of the 110 effect sizes could not be calculated based on the means and standard deviations of posttest scores. The way the effect size was calculated (posttest scores or another method) was found not to be a significant moderator ($p = .653$).

[2] For SRL →achievement 17 of the 172 effect sizes could not be incorporated based on Pearson's *r*. The effect size used (Pearson's *r* or another correlational measure) was found not to be a significant moderator ($p = .915$).

Table 3
Sample data for MASEM preprocessing.

ID	Study	Relation	*n*	ES
1	1	1	80	0.6
2	1	1	80	0.5
3	1	2	70	0.5
4	1	2	60	0.5
5	1	3	90	0.4

estimates as the weight matrix is used to estimate the path model coefficients.

To conduct MASEM all effect sizes must be expressed in Pearson's *r*. The effect sizes for the effect of SRL interventions on achievement and SRL activities, which were expressed in Cohen's *d*, were converted using the formula $r = d/(\sqrt{d^2 + a})$ with correction factor $a = (n_1 + n_2)^2/(n_1 n_2)$ (Borenstein, Hedges, Higgins, & Rothstein, 2009). It is not yet possible to conduct three-level MASEM for nested data. Therefore, if a study contained multiple effect sizes for the same relation, these had to be combined. This process is explained with the use of sample data presented in Table 3.

First, the data was merged to make sure there was one effect size with one sample size per relation per study. In the sample data there are two effect sizes for relation 1 in study 1: ID 1 and 2. These studies have the same sample size, but a different effect size. These effect sizes must be merged. A weighted average of these effect sizes is taken based on *n*. The effect size for relation 1 in study 1 is therefore 0.55 with $n = 80$. The effect sizes for relation 2 in study 1 must also be merged (ID 4 and 5). The weighted effect size is 0.5. For each relation however, only one sample size can be included. Because for relation 2 in Table 3, these sample sizes differ, the average sample size is used. In this case this is 65. Second, the data had to be merged further to make sure there was one sample size per study with up to three effect sizes (one effect size per relation). The sample size is the harmonic mean of the three sample sizes (80, 65, and 90). Leading to the following data for study 1: $n = 76.93$, ES relation 1 = 0.55, ES relation 2 = 0.5, and ES relation 3 = 0.4.

The mediation MASEM model allows us to evaluate the significance and size of the indirect effect of SRL interventions on achievement through SRL activity. If, in addition to a significant indirect effect, there is a direct effect of SRL interventions on achievement, it would indicate that the effect is *partially* mediated by SRL.

Although MASEM is the analytical technique to evaluate indirect effects between variables, the technique also has its shortcomings. Specifically, it is not (yet) possible to apply three-level models to account for nested effect sizes within studies in MASEM, or to evaluate the effect of continuous moderator variables. In our dataset, a large number of studies provided multiple effect sizes for the same relation. For example, a study may measure multiple aspects of SRL, and correlate each of them separately to students' grades. These so-called nested effect sizes within studies can be controlled for with multilevel analyses (Hox, 2010). Therefore, we also conducted univariate analyses using a meta-analytic method for handling complex meta-analytic data structures with robust variance estimation. Robust variance estimation accounts for the dependency of effect sizes within studies (Hedges, Tipton, & Johnson, 2010).

The univariate meta-analyses for the effect of SRL interventions on achievement and SRL (relation 1 and 2) were conducted with Cohen's *d* as a measure of effect size. The analyses for the relation between SRL and achievement (relation 3) were based on Pearson's *r*. As the variance in Pearson's *r* however depends too strongly on the correlation, the analyses were conducted in Fisher's Z (Borenstein, , Hedges, , Higgins, , & Rothstein, 2009). Pearson's *r* was converted into Fisher's Z with the following formula: $z = 0.5 * \ln((1 + r)/(1 - r))$. Then, the results were converted back to Pearson's *r* with the formula $r = (e^{2z} - 1)/(e^{2z} + 1)$. All results for the analyses concerning the relation between SRL and achievement are reported in Pearson's *r* to simplify interpretation, but were conducted in Fisher's Z.

As we did not expect a single true effect of SRL interventions on either SRL or achievement, nor a single true effect size of the relation between SRL and achievement, the MASEM analysis and all meta-analyses were conducted based on the random-effects model (Borenstein et al., 2009). Hereby we allowed the effects to vary from study to study, and thus accounted for heterogeneity of effect sizes.

The analyses on publication bias were conducted using CMA v3.3 (Borenstein, Hedges, Higgins, & Rothstein, 2014). The MASEM analysis was conducted using the metaSEM package (Cheung, 2015) for R (R Core Team, 2016) and parts of the example script created by Jak (2015). The univariate multilevel meta-analyses and moderator analyses were conducted using the Robumeta package (Fisher, Tipton, & Zhipeng, 2017) for R (R Core Team, 2016). This package implements the methods developed by Hedges, Tipton, and Johnsen (2010) for handling complex meta-analytic data structures with dependent effect sizes using robust variance estimation. The data files and syntax for both the MASEM analyses and the univariate moderator analyses are available through a Mendeley data repository linked to this article.

3. Results

3.1. Descriptive statistics

A total of 395 effect sizes extracted from 142 studies published in 126 articles were included in the meta-analyses. The average sample size of the included effect sizes was $n = 214$ (range 13–8112). From the 126 articles, 112 originated from peer-reviewed

journals, 9 were published in conference proceedings, and 5 were dissertations. In the majority of studies (117 out of 142 studies), participants were undergraduates. In other studies participants were in graduate education (5 studies), in vocational education (4 studies), in workplace learning (3 studies), or their educational level could not be determined (13 studies).

Some articles contained data from multiple studies, and some studies contained multiple effect sizes on the same relationship, or effect sizes on multiple relationships. For the effect of SRL interventions on achievement (relation 1), 112 effect sizes from 61 studies published in 51 articles were included. For the effect of SRL interventions on SRL activity (relation 2), 111 effect sizes were included extracted from 41 studies published in 32 articles. Finally, for the relationship between SRL and achievement (relation 3), 172 effect sizes were included from 72 studies published in 70 articles. An overview of all included studies can be found in Supplementary Table S1 (online only). Tables S2, S3, and S4 (online only) present overviews of the effect sizes included for relation 1, relation 2, and relation 3 respectively.

Potential outliers were identified by creating boxplots. As all included effect sizes were calculated from published primary studies, there had to be a compelling reason to consider an effect size an outlier. One effect size was determined to be an outlier (Cohen's $d = 5.7733$) as it was more than two times the size of the second largest effect size (Cohen's $d = 2.5192$). In the experiment that yielded the largest effect size, students were requested to watch a lecture and afterwards to write about their learning process while watching the lecture. Students in the experimental group were prompted to write about their comprehension monitoring and strategy regulation. The researchers afterwards counted the number of times students wrote about these topics (Berthold, Nückles, & Renkl, 2007). The prompting led to an inflated count for the students in the experimental group. This effect size, for the relation between SRL interventions and SRL, was excluded from further analysis. Therefore, for the effect of SRL interventions on SRL not 111, but 110 effect sizes were included from 41 studies published in 32 articles.

3.2. Publication bias

Studies with larger samples and studies with significant results are more likely to be published than studies with smaller samples and studies with non-significant results (Borenstein et al., 2009). This results in publication bias, which may lead to bias in the sample set selected for inclusion in a meta-analysis. As there is no formal statistical test to determine the presence of publication bias and its effect, it is advised to combine several analyses (Banks, Kepes, & Banks, 2012; Borenstein et al., 2009). In the current study, the funnel plot with trim and fill, cumulative meta-analysis, Orwin's fail safe N (Orwin, 1983), adding the sample size as a moderator, and adding the source of the study (journal article, conference proceedings, or dissertation) were used to test for the potential presence of publication bias in the sample sets included in these meta-analyses (Banks et al., 2012; Borenstein et al., 2009; Hox, 2010). The analyses were conducted for each of the three relations in Fig. 1 separately. From the results of these five tests for publication bias, it can be concluded that there are no indications that publication bias has affected either of the three sets of included effect sizes. Source of publication was however found to be a significant moderator of the effect sizes for relation 1 (intervention – achievement) and relation 2 (intervention – SRL). Source of publication will therefore be included as a covariate in all moderator analyses conducted for these relations. A more detailed description of the publication bias analyses and their results is available as supplementary material (online only).

3.3. MASEM analysis

In Stage 1 of the MASEM analysis, a pooled correlation matrix between the three variables was estimated. All three pooled correlations were positive, and significantly larger than zero ($r_{\_int\_achievement} = 0.22$, $p < .05$; $r_{\_int\_SRL} = 0.22$, $p < .05$; $r_{\_SRL\_achievement} = 0.25$, $p < .05$), indicating medium sized correlations for all three relations (Cohen, 1988). These average correlation coefficients from the Stage 1 TSSEM analysis, together with the associated number of studies and total sample sizes are presented in Supplementary Table S5 (online only). Next, we fitted the mediation model from Fig. 1 to the pooled correlation matrix. The estimates of the indirect (relation 2 and 3 combined) and direct effects (relations 1, 2, and 3) can be found in Table 4. A significant indirect effect was found of SRL interventions on achievement. The indirect effect equals the product of the direct effect of SRL interventions on SRL ($\beta = 0.22$, $p < .05$) and the direct effect of SRL on achievement ($\beta = 0.22$, $p < .05$). The significant indirect effect indicates that part of the effect of SRL interventions on achievement is mediated by SRL activity, although the indirect effect is small ($\beta = 0.05$, $p < .05$). Furthermore, a significant direct effect of interventions on achievement was found ($\beta = 0.18$, $p < .05$). This existence of the significant direct effect while incorporating SRL activity as a mediator, indicates that the mediational effect through SRL is partial as opposed to full.

Table 4
Standardized regression coefficients of the stage 2 MASEM analysis.

Relation	β	95% CI
SRL interventions → achievement	0.18	[0.13; 0.22]
SRL interventions → SRL	0.22	[0.16; 0.28]
SRL → achievement	0.22	[0.16; 0.27]
SRL interventions → SRL → achievement	0.05	[0.03; 0.06]

Table 5
Results of random-effects univariate meta-analyses with robumeta.

Relation	k	# of ES	Intercept (std. Error)	τ^2	I^2
SRL interventions → achievement	61	112	d = 0.488 (.053)	0.126	68.95
SRL interventions → SRL activity	41	110	d = 0.499 (.080)	0.206	78.61
SRL activity → achievement	72	172	z = 0.284 (.031) r = 0.277	0.050	94.43

Note. τ^2 is the between-study variance component. I^2 represents the proportion of variance in effect sizes at the study-level. r is the z-to-r back-transformed correlation coefficient.

3.4. Univariate multilevel meta-analyses

The results of the univariate multilevel meta-analyses are presented in Table 5 and indicate medium sized effects for all three relations; a medium sized effect of SRL interventions on achievement, a medium sized effect of SRL interventions on SRL and a medium sized relation between SRL and achievement.

3.5. Moderator analyses

While the overall effect sizes are found to be significant (Table 5), the effect sizes within each set of studies vary as indicated by the I^2 values, which are above 65% for all three relations (Higgins & Thompson, 2002). Moderator analyses were conducted to determine if aspects of the included studies can explain part of the variance in effect sizes. Three groups of moderators were coded: study characteristics, measurement characteristics, and intervention characteristics. Moderator analyses were conducted for each relation separately, and each group of moderators was tested separately for significance. By testing the moderators per group, we balance the risk of type I errors (that would be induced by testing all moderators individually) and type II errors (that would be induced by testing all moderators simultaneously).

3.5.1. Moderator analyses SRL interventions → achievement

For the effect of SRL interventions on achievement, moderator analyses were conducted with all study and intervention characteristics. Furthermore, the employed achievement measure was included as a moderating measurement characteristic. The result of the moderator analyses are presented in the left columns of Tables 6–8. Study characteristics (left columns Table 6) did not moderate on the effect of SRL interventions on achievement. The only exception was academic subject, with higher effect sizes for studies conducted within the humanities domain compared to social sciences. Effect sizes were found to differ based on the achievement

Table 6
Results of the moderator analyses with study characteristics.

Moderator	Intervention → achievement					Intervention → SRL					SRL → achievement				
	k	#ES	Est	SE	p	k	#ES	Est	SE	p	k	#ES	Est	SE	p
Intercept	61	112	0.425	.105	.001	41	110	0.650	.210	.011	72	172	0.306	.064	.000
Academic subject															
Social sciences (ref)	17	41				15	34				26	46			
Humanities	10	12	0.607	.251	.026	4	20	−0.111	.323	.750[+]	3	8	0.060	.076	.479[+]
Formal sciences	14	22	0.100	.100	.324	6	13	−0.129	.256	.625	11	27	0.138	.102	.192
Applied sciences	13	23	0.070	.115	.548	4	8	0.175	.391	.671	18	57	−0.160	.090	.079
Mixed	7	14	0.011	.163	.948	11	31	−0.266	.153	.113	12	30	−0.084	.090	.362
Setting															
Study offline (ref)	29	51				26	75				61	144			
Study online	32	61	−0.191	.130	.157	15	35	−0.254	.301	.420	11	28	0.065	.139	.645
Quality															
Quality level 1 (ref)	40	63				20	45								
Quality level 2	16	38	−0.039	.109	.727	12	49	−0.241	.207	.266					
Quality level 3	5	11	0.110	.209	.616	9	16	0.001	.192	.997					
Context															
Real course (ref)	35	61				29	86				64	151			
Lab	26	51	0.138	.126	.285	12	24	0.342	.345	.350	8	21	−0.077	.064	.261

Note. k = number of studies; #ES = number of effect sizes.
For intervention → achievement and intervention → SRL, source was a significant moderator. Therefore source is included as a covariate in the moderator analyses conducted for these relations.
For intervention → SRL, there is 1 study with 4 effect sizes for which information on academic subject is missing. The analysis was also conducted without academic subject as a moderator. This did not alter the significance of the results.
For SRL → achievement, there are 2 studies with in total 4 effect sizes for which information on academic subject is missing. The analysis was also conducted without academic subject as a moderator. This did not alter the significance of the results.
[+] df < 4; the result should not be interpreted.

Table 7
Results of the moderator analyses with measurement characteristics.

Moderator	Intervention → achievement				Intervention → SRL				SRL → Achievement			
	#ES	Est	SE	*p*	#ES	Est	SE	*p*	#ES	Est	SE	*p*
Intercept	112	0.605	.109	.000	110	0.521	.055	.000	172	0.334	.081	.002
Instrument SRL												
Questionnaire (ref)									147			
Count					16	0.825	.289	.014	10	0.095	.142	.539[+]
Self-assessment accuracy					12	−0.056	.129	.672	15	0.224	.128	.094
SRL activity measured												
General (ref)									13			
Metacognitive multiple aspects					18	−0.169	.140	.247	35	−0.155	.090	.103
Metacognitive single aspect					59	−0.201	.130	.147	51	−0.083	.111	.459
Resource management multiple aspects					8	0.028	.214	.900	13	−0.213	.087	.028
Resource management single aspect					18	−0.285	.122	.039	60	−0.214	.089	.026
Achievement measure												
Course performance (ref)	54								107			
Task performance	46	−0.157	.117	.185					20	−0.013	.074	.867
GPA	1	−0.241	.109	.037					41	0.091	.058	.122
Transfer task	11	−0.206	.154	.224					3	−0.012	.045	.807[+]

Note. #ES = number of effect sizes.
For intervention → achievement and intervention → SRL, source was a significant moderator. Therefore source is included as a covariate in the moderator analyses conducted for these relations.
For SRL → achievement, there is 1 study with 1 effect size for which information on the achievement measure used is missing. The analysis was also conducted without achievement measure as a moderator. This did not alter the significance of the results.
[+] df < 4; the result should not be interpreted.

Table 8
Results of the moderator analyses with intervention characteristics.

Moderator	Intervention → achievement					Intervention → SRL				
	k	#ES	Est	SE	*p*	k	#ES	Est	SE	*p*
Intercept	61	112	.728	.131	.000	41	110	0.355	.258	.212
SRL activity supported										
General (ref)	19	30				16	40			
Metacognitive multiple aspects	25	52	−0.238	.170	.174	16	50	0.286	.246	.279
Metacognitive single aspect	17	30	−0.186	.150	.227	9	20	−0.017	.178	.924
Exclusively SRL										
Exclusively SRL (ref)	41	70				26	72			
Not exclusively SRL	20	42	0.169	.115	.156	15	38	−0.012	.233	.961
Timing										
During whole course (ref)	42	71								
Before the course	1	6	−0.298	.193	.309[+]	30	77			
First half of course	3	6	−0.165	.160	.390[+]	3	13	−0.107	.434	.822[+]
Second half of course	7	8	−0.013	.163	.937	2	6	0.890	.707	.327[+]
Before and during course	8	21	−0.215	.128	.115	6	14	−0.031	.328	.927
Tailored										
Tailored to context (ref)	60	110				34	98			
Not tailored to context	1	2	−0.467	.263	.156[+]	7	12	−0.005	.249	.984
Format										
Instruction	27	60	0.086	.103	.412	21	53	0.152	.264	.574
Application	43	85	−0.042	.092	.656	30	86	−0.018	.213	.935
Prompting	28	51	−0.123	.116	.302	17	37	−0.121	.285	.678

Note. k = number of studies; #ES = number of effect sizes.
For intervention → achievement and intervention → SRL, source was a significant moderator. Therefore source is included as a covariate in the moderator analyses conducted for these relations.
[+] df < 4; the result should not be interpreted.

measure used in the experiment, as studies that used GPA as an indicator of academic achievement reported lower effect sizes compared to studies that used course performance as an indicator of academic achievement. As this result is however based on only 1 effect size with GPA as achievement measure, it should not be interpreted (left columns of Table 7). None of the intervention characteristics significantly moderated the effect sizes (left columns Table 8).

3.5.2. Moderator analyses SRL interventions → SRL

For the effect of SRL interventions on SRL activity, moderator analyses were conducted with all study and intervention characteristics. From the measurement characteristics, the instrument used to measure SRL and the measured SRL activities, were included as potential moderators. The results of the moderator analyses are presented in the middle columns of Tables 6 and 7 and the right columns of Table 8. None of the included study characteristics significantly moderated the effects of SRL interventions on SRL (middle columns of Table 6). The instrument used to measure SRL was a significant moderator, with larger effects of SRL interventions on SRL activity when SRL activity was counted compared to when SRL activity was measured with a questionnaire. The measured SRL activity was also a significant moderator; the measurement of a single resource management aspect resulted in smaller effect sizes compared to the measurement of both resource management and metacognitive aspects of SRL (middle columns of Table 7). None of the intervention characteristics were significant moderators of the effects of SRL interventions on SRL (right columns of Table 8).

3.5.3. Moderator analyses SRL → achievement

For the relationship between SRL and achievement, moderator analyses were conducted with all mentioned study and measurement characteristics. As these studies did not include an intervention, no moderator analyses were conducted with intervention characteristics. The results of the moderator analyses are presented in the right columns of Tables 6 and 7 None of the study characteristics significantly moderated the relationship between SRL and achievement (right columns Table 6). The instrument used to measure SRL was also not a significant moderator. The measured SRL activities were significant moderators. The relationship between SRL and achievement was stronger when both metacognitive and resource management activities were measured compared to when a single or multiple resource management aspects were measured (right columns Table 7).

4. Discussion

The aim of our meta-analysis of SRL in higher education was twofold. The first aim was to study the potential mediating role of SRL activity in the relationship between SRL interventions and academic achievement. The second aim was to perform meta-analyses of the effect of SRL interventions on achievement (relation 1 in Fig. 1), the effect of SRL interventions on SRL activity (relation 2 in Fig. 1), and of the relationship between SRL activity and achievement (relation 3 in Fig. 1). For each of these meta-analyses, moderator analyses were conducted to determine which study, measurement, and intervention characteristics could significantly explain variability in the effect sizes reported. The results of the moderator analyses may help develop guidelines for the design of effective SRL interventions.

4.1. Mediation of SRL activity in the relationship between SRL interventions and academic achievement

SRL activity was expected to mediate the relationship between SRL interventions and achievement: the SRL interventions were assumed to lead to improvements in students' engagement in SRL, and thereby to improvements in students' achievement. The result of the MASEM analysis confirmed that SRL activity mediates the effect of SRL interventions on achievement. However, the results also show that this indirect effect of SRL interventions on achievement is small ($\beta = .05$), and that a significant direct effect of SRL interventions on achievement remains after including SRL activity as a mediator ($\beta = 0.18$). Contrary to common belief, SRL activity is thus only a *partial* mediator of the effect of SRL interventions on achievement. So, on the one hand, the results support the often made assumption that improvement in student achievement after implementing an SRL intervention is due to improvements in students' SRL activities. On the other hand, the results also indicate that there are other factors leading to improvements in student achievement resulting from an SRL intervention, besides improvements in SRL activity. We therefore conclude that SRL interventions likely have side effects on other factors than SRL activity, which in turn lead to improvements in achievement.

4.2. Explanations for SRL activity as partial mediator

Several other factors potentially influencing the effect of SRL interventions on achievement can be identified. A possible explanation for the partial mediation of the effect of SRL interventions on achievement by SRL activity are difficulties with SRL measurements. Several authors have argued that self-report measures of SRL, which are used in a large number of the studies included in this review, are not always valid and reliable (Azevedo & Cromley, 2004; Hadwin, Nesbit, Jamieson-Noel, Code, & Winne, 2007; Winne & Perry, 2000). SRL is a process, while questionnaires do not measure how students strategically adapt their study tactics over time. Furthermore, the validity of questionnaire measures of SRL is debated (Veenman, 2016; Winne, Jamieson-Noel, & Muis, 2002). If the employed SRL measurements were insensitive to changes in students' engagement in SRL activities over time, then improvements in SRL activity due to the SRL intervention would not be measured accurately. This insensitivity would lead to an underestimation of the mediating role of SRL activities. More precise measures of SRL that are better able to accurately capture small changes in students' SRL activities could allow for an even better estimation of the mediating role of SRL activities in the relationship between SRL interventions and achievement. We refer the interested reader to the work of Panadero et al. (2016), Veenman (2016; Veenman et al., 2006), and Winne (2010; Winne et al., 2002) for an in-depth discussion of the issues concerning SRL measurement.

Another potential explanation for the partial mediation is the possible side effect of SRL interventions that they may not only lead to changes in students' engagement in SRL activities, but also to increased time on task for students. Stimulating students to monitor or reflect on their learning can trigger students to engage in corrective action, thereby making them spend additional time on learning

that they would not have spent without the intervention (Belski & Belski, 2014). If time on task differed between intervention and control groups, then this could explain differences in academic achievement. Unfortunately, student time on task is not measured (or reported) in any of the experimental SRL intervention studies included in our analyses.

SRL interventions may not only lead to increased time on task, but also to an increased use of more effective cognitive activities during learning, as students' increased self-regulation may have caused them to engage in more, or more effective, cognitive activities. As described by Nelson and Narens (1990), learners have a mental representation of their learning process at a meta-level. Based on the discrepancy between this meta-level representation and the goal level of learning, learners regulate their learning at the object-level. Learners, for example, engage in different cognitive learning activities. Learners then monitor their learning at the object-level as feedback for the meta-level (Nelson & Narens, 1990). The inclusion of cognitive strategy training did not significantly moderate the effects of SRL interventions on either SRL activity or on achievement. However, an improvement in self-regulation activities may have caused students to engage in more effective cognitive activities. For instance, repeated monitoring of the learning progress may have caused students to realize what cognitive activities do and do not benefit their learning. The increased engagement in effective cognitive strategies would then likely lead to increased student achievement (Dunlosky et al., 2013). In the current review, we focused on the effects of SRL interventions on achievement, engagement in cognitive activities therefore lies beyond the scope of the current review. Cognitive activities may, however, have explained part of the direct effect of SRL interventions on achievement. The role of cognitive strategies in the effects of SRL interventions on achievement is therefore an interesting topic for a future review study.

Finally, task motivation (i.e., motivation before learning) is described as consisting of self-efficacy, task value, and goal orientation. Task motivation also lies beyond the scope of the current review, as we considered it a precursor of SRL (cf. Artino & Stephens, 2006; Efklides, 2011; Ning & Downing, 2012; Pintrich, 1999). However, as described in the Introduction, task motivation is considered part of SRL in some theoretical models (e.g., Boekaerts, 1992; Zimmerman, 2002, 2008). If we would have included task motivation in our definition of SRL, then task motivation measures such as self-efficacy and goal orientation, would have been included in the analyses. This would have broadened the construct 'SRL activity' in Fig. 1. Thereby, SRL activity would most likely have explained a greater portion of the effect of SRL interventions on achievement, thus strengthening the mediation effect. The strength of the mediation by task motivation of the effect of SRL intervention on achievement is a worthwhile research question to gain further insight in the interrelations between SRL interventions, SRL activity, task motivation, and achievement.

4.3. Meta-analyses of the relations between SRL interventions, SRL activity, and achievement

The second aim of this study was to systematically review the effect of SRL interventions on achievement, the effect of SRL interventions on SRL activity, and the relationship between SRL activity and achievement in higher education. We found medium effect sizes for all three relations. SRL interventions are thus effective for supporting students' achievement and students' engagement in SRL activities.

The sizes of the effects of SRL interventions on achievement and on SRL activities are smaller than the effect sizes reported in previous meta-analyses in primary and secondary education (Boer et al., 2013; Dignath & Büttner, 2008; Dignath et al., 2008). Students in higher education may have internalized a larger repertoire of SRL strategies than pupils in primary and secondary education. Dignath and Büttner (2008) already found that pupils in secondary education benefitted less from SRL instruction compared to learners in primary education. These authors argued that pupils in secondary education likely had internalized a larger repertoire of SRL strategies and thus benefitted more from prompts to make use of these strategies. The smaller effect of SRL instruction on achievement found in this study compared to reviews conducted in primary and secondary education may be explained in a similar manner: Instructing students on SRL may be less effective in higher education as there is less room for improvement. Further research to substantiate this claim is however necessary, preferably by reviewing the (difference in) effects of SRL interventions across educational levels.

The expertise reversal effect may present an additional explanation for SRL interventions being less effective in higher education compared to primary and secondary education. The expertise reversal effect refers to the interaction between the effectiveness of support and students' level of expertise: while students with low abilities benefit from support, students with high abilities are impeded by support (Schnotz, 2010). If there are indeed students in higher education that possess an adequate repertoire of SRL strategies, then the tasks presented in some of the studies may not have posed a sufficient SRL challenge for these students. The performance of high ability students in these tasks would then have been hampered by offering SRL support, as the support would have distracted them from learning (Clarebout, Horz, Schnotz, & Elen, 2010; Pieger & Bannert, 2018). The overall effect sizes would thereby be lowered. Further research including measurement of learners' expertise could determine whether the expertise reversal effect indeed occurs with SRL interventions in higher education. For more information on the expertise reversal effect, we refer the interested reader to the work of Van Merriënboer and others (Van Merriënboer & Sweller, 2010; Young, Van Merriënboer, Durning, & Ten Cate, 2014).

While the need for SRL support may be reduced in higher education, the need for engagement in SRL activities increases in higher education due to the increase in student autonomy. The increased autonomy would imply that the relation between engagement in SRL activities and achievement would be stronger in higher education than in primary and secondary education. When comparing the relation between engagement in SRL activities and achievement for higher education in our meta-analysis ($r = 0.28$), with the relation reported by Dent and Koenka (2016) for primary and secondary education ($r = 0.20$), it becomes clear that engagement in SRL activity and achievement are more strongly related in higher education than in primary and secondary education. The correlation we found is also in line with the previous meta-analysis in higher education that examined the relation between SRL or specific

SRL activities and achievement (Sitzmann & Ely, 2011). Correlations in this meta-analysis were found to range between $r = 0.11$ and $r = 0.37$.

The effect sizes we found were not homogeneous, but heterogeneous; their magnitude varied between studies. Moderator analyses were conducted to test which study, measurement, or intervention characteristics could significantly explain (part of) the variance in effect sizes between studies (see Table 2).

4.4. Measurement characteristics tested as moderators

We first discuss the measurement characteristics that were included as moderators. The instrument used to measure SRL did not significantly influence the effect of SRL activity on achievement. Interventions did have significantly different effects on SRL activity depending on how SRL activity was measured; effects were stronger when SRL activity was measured with a counted measure, compared to when SRL activity was measured with a questionnaire. The differences in effect sizes between studies with different measurement instruments may have been the result of measurement (in)validity. Questionnaire measures of SRL are debated, since students' self-reports may be inaccurate and questionnaire measures are not able to capture learners' SRL over time (Hadwin et al., 2007; Veenman, 2016). Learners' self-reported SRL has been found to be weakly correlated to actual counts of strategy use (Hadwin et al., 2007; Veenman, 2005). While this explanation is plausible, yet another explanation focuses on the setting in which the studies were conducted. Counted measures were mostly used in studies in which participants were stimulated to think aloud (e.g., Moos & Bonde, 2016), and their verbalizations were coded for utterances of SRL. All studies with counted measures were conducted in a lab setting, while almost all studies with questionnaire measures (77 of 82 effect sizes) were conducted in real courses. Interventions conducted in lab settings usually have greater intensity than interventions implemented in regular education, especially when participants are required to think aloud, as the experimenter is then continuously present while the participant studies the presented material. The experimenter not only prompts the student to verbalize his/her thoughts during learning, but also – implicitly – makes sure that the student adheres to the intervention. The increased intensity of lab studies using counted measures, and the resulting intervention fidelity, may thus explain why counted measures have resulted in larger effect sizes compared to questionnaire measures. Further research is necessary to determine what causes the increased effect sizes of SRL interventions on SRL activity when SRL is measured with a questionnaire (i.e., an offline measure) compared to a counted measure (i.e., an online measure).

The effect of SRL interventions on SRL activity was found to be stronger for general SRL measurements that included both metacognitive and resource management activities, compared to measures that focused on a single aspect of resource management (e.g., time management). While the effects of SRL interventions on resource management are often measured, there are no SRL interventions that solely focus on resource management; interventions either focus on metacognition and resource management, or only on metacognition. It is therefore not surprising that the effects of the interventions are smaller on resource management than on SRL in general, as resource management is either not at all part of the intervention (metacognitive interventions) or only a part of the intervention (general interventions). The difference in effect sizes is thus likely due to the misalignment of SRL interventions and SRL measures concerning resource management.

Besides moderating the effect of SRL interventions on SRL activity, the measured SRL activity was also found to be a significant moderator of the relationship between engagement in SRL activities and achievement. General measures of SRL (i.e., including both metacognition and resource management activities) were more strongly related to achievement than both single and multiple aspect measures of resource management aspects of SRL. Metacognitive measures did not differ significantly from general measures, but the coefficients indicate that the strength of their relationship to achievement was in between those for general and resource management measures. Scoring highly on general SRL measures indicates that the student engages in a range of SRL activities, both metacognitive and resource management activities, instead of only resource management activities. Likely, this greater diversity also indicates greater engagement with SRL, especially as resource management activities are used only during the performance phase of learning, while metacognitive activities are used also before learning (planning) and after learning (reflection). As a result, general SRL measures were more strongly associated with achievement than resource management measures of SRL.

The final measurement characteristic that was tested as a moderator was the achievement measure used in the included studies. Almost all of the included studies used task performance, course performance, or GPA as achievement measure. A greater similarity between the training task and the achievement measure might have resulted in a greater effect size (Burke & Hutchins, 2007). The overlap between training and test is usually greater for task performance measures compared to course performance or GPA. Therefore, effect sizes might have been larger with task performance compared to course performance or GPA. However, no significant differences were found in the effect sizes. In the current study, achievement measure thus did not influence the effect of SRL interventions on achievement, nor the relationship between SRL and achievement.

4.5. Intervention characteristics as moderators

We now turn to the intervention characteristics that were tested as moderators of the effect of SRL interventions on either engagement in SRL activities or on achievement. Contrary to expectations, none of the intervention characteristics were found to significantly moderate either the effect of SRL interventions on SRL activity, or the effect of SRL interventions on achievement.

All intervention formats (instruction, format, prompting) and possible timings of the intervention (before the course, first half or second half of course, before and during the course, and during the whole course) were found to be beneficial for both SRL activity as well as achievement; none had a negative effect. However, no significant differences between the intervention formats and timings of the interventions were found. The inclusion of cognitive strategies in the SRL intervention was also an insignificant moderator on

both relations. This finding is in line with the assumption that cognitive strategy training is important for younger children, but not for students in higher education. In contrast to students in higher education, young children have not yet internalized a sufficient repertoire of cognitive strategies they can implement during learning (Dignath & Büttner, 2008). No significant differences were found between interventions tailored to the learning context and interventions that were not tailored to the learning context either. The group of studies with interventions tailored to the learning context was, however, much larger than the group of studies with interventions that were not tailored to the learning context. Therefore, we cannot determine whether there is no influence of SRL interventions being tailored to the learning context, or if the lack of a significant result is due to too few studies with a non-tailored intervention design. The results of the analyses with the format of the intervention (i.e., instruction, application or prompting) as moderator, however, also do not provide clear indications that tailoring the intervention to the learning context is crucial for the effectiveness of the SRL intervention to support students' engagement in SRL activities and/or their achievement.

The final intervention characteristic we tested was the SRL activity or set of SRL activities supported by the SRL intervention. Based on the finding that general SRL measures were more strongly related to achievement than resource management measures, it was anticipated that general SRL interventions would be more beneficial for achievement than interventions that focus on either metacognition or resource management. Developing a broad set of SRL skills could be considered more important for achievement. However, no differences were found in SRL intervention effectiveness depending on the set of SRL activities supported.

4.6. Non-significance of moderators

From the large range of moderators tested, only a few were found to be significant. Multicollinearity between coded characteristics may explain the lack of significant results. Some study, measurement, and intervention characteristics occurred together frequently, such as prompting in lab studies and using questionnaires in real courses. To reduce the risk of not finding significant results due to multicollinearity (type II errors), we tested each group of characteristics separately. However, there also was some multicollinearity between characteristics within each group. For instance, prompting only occurred during studies and not before. Furthermore, the SRL measurement and the type of SRL measured are related. Almost all measures of multiple SRL aspects (i.e., general, metacognitive multiple aspects, and resource management multiple aspects) were measured with questionnaires, and not with self-assessment accuracy or counted (e.g., think aloud) measures. Self-assessment accuracy on the other hand is a measure of self-assessment which is always coded as a 'single aspect metacognitive measure'. All moderators of one group were tested at once, as testing all moderators separately would have greatly increased the number of analyses and thereby the risk of falsely finding significant results. However, the frequent co-occurrence of specific characteristics within groups may have reduced the relative predictive strength of each moderator.

In addition, three aspects of the included studies may have significantly influenced the relations tested. These aspects may therefore have been significant moderators. They were however not included as moderators in the analyses as they could not be coded. First, as already indicated in the discussion of the mediation results, time on task may have differed between conditions. If the intervention caused students to spend more time learning, then time on task may have moderated the effect of SRL interventions on achievement. Time on task could however not be coded as it was not reported in any of the included studies. Another potential moderator that could not be coded was the intensity of the intervention, while it could be assumed that interventions with a greater intensity likely have greater effects. For instance, a 30-min instruction on SRL right before a 45-min learning task (Azevedo & Cromley, 2004) likely has a much greater effect compared to having students work through 10 modules of materials and the intervention consisting of a self-reflection prompt ("Do I understand all of the key points of the training material?") after five of these modules (Sitzmann, Bell, Kraiger, & Kanar, 2009). It was attempted to code intervention characteristics that indicate intervention intensity (e.g., weeks the intervention ran, duration of the intervention in hours). In most studies however, this information was not reported clearly enough to be coded. Therefore the intervention intensity could not be tested as a moderator. A third moderator that could not be included due to missing information was participants' and teachers' fidelity to the intervention (e.g., the extent to which teachers instructed SRL as intended by the researchers). Participants and teachers likely adhered to the interventions in different degrees, which could influence the effects of the SRL intervention on both SRL activity and achievement. We consider these three potential moderators important to take into account in future SRL studies, as described as suggestions for future research.

Finally, students' motivation may also have influenced the tested relations. Students' with greater task motivation (e.g., self-efficacy) engage in SRL activities more frequently (Artino, 2008; Pintrich, 1999). Task motivation may also influence the effectiveness of SRL interventions, with students who are more interested in the task, also being more interested in learning ways to improve their SRL and achievement. Interventions would then be more effective in improving engagement in SRL activities and achievement for students with greater task motivation. Task motivation may therefore moderate the effect of SRL interventions on SRL activities as well as the effect of SRL interventions on achievement. In the current review, task motivation was not included as it is considered a precursor of SRL (Artino & Stephens, 2006; Efklides, 2011; Ning & Downing, 2012; Pintrich, 1999). However, as already indicated in the discussion of the mediation effects, the influence of task motivation on the relations between SRL interventions, SRL activities, and achievement is a worthwhile direction for further review studies to explore.

4.7. Implications and future research

To summarize, we investigated the mediating role of SRL activity on the effect of SRL interventions on achievement, and we explored which moderators significantly explain the variability in effect sizes for (1) the effect of SRL interventions on achievement, (2) the effect of SRL interventions on SRL activity, and (3) the relation between SRL activity and achievement. In this section, we

discuss practical, theoretical, and methodological implications of our findings and suggestions for future research.

The positive effect of SRL interventions on both SRL activity as well as on achievement, leads to the practical implication that SRL interventions are effective in supporting students. We therefore advice practitioners to implement SRL interventions in higher education to support learners' engagement in SRL activities as well as their achievement. The lack of significant moderators of the effects of SRL interventions makes it difficult to provide concrete design guidelines for such SRL interventions. However, the lack of significant moderators also implies that there may be no 'wrong' intervention designs; as no intervention type is less effective than others, practitioners have great freedom in their intervention design. Based on our results, we thus cannot offer practical guidelines on *how* to develop the intervention, but we *do* emphasize the importance of implementing an SRL intervention to support learners' engagement in SRL activities and their achievement.

The partial mediation of SRL interventions on achievement by SRL activity results in the theoretical implication that the improvements in student achievement as a result of SRL interventions are *mostly* due to other factors than students' SRL activity. More insight should be gained on what these *additional factors* are or could be. We have described the influence time on task, cognitive activity, and task motivation may have on the effectiveness of SRL interventions. It might be worthwhile to review the relations between SRL interventions, SRL activities, achievement, time on task, cognitive activity, and task motivation to further explain the effect of SRL interventions on achievement.

The differential effects of SRL interventions on SRL and achievement lead to methodological implications. As the results of our mediation analysis indicate, it cannot be assumed that the effect of SRL interventions on achievement is (solely) due to students' engagement in SRL activities. Therefore, it is important to test and report the effects of SRL interventions on both SRL activity and achievement, and that the correlation between SRL activity and achievement is reported as well. Doing so allows for the calculation of the mediating effect of SRL activity in the relation between SRL interventions and achievement, which is important for our understanding of the relation between these constructs, as well as for the design of SRL interventions.

Finally, several moderators could not be tested due to information not being (clearly) reported in the studies included in this meta-analysis (i.e., time on task, intervention intensity, and intervention fidelity). These moderators may however significantly explain part of the variance in effect sizes. We therefore recommend researchers to report information on these aspects in future studies, making it possible to test and review their influence.

5. Conclusion

In this study, we have performed meta-analyses of three relations: (1) the effect of SRL interventions on achievement, (2) the effect of SRL interventions on engagement in SRL activities, and (3) the relationship between engagement in SRL activities and achievement. With advanced statistical methods, we tested whether improvements in achievement after implementing an SRL intervention are mediated by SRL activity. We found evidence that the effect of SRL interventions on achievement is only partially mediated by SRL activity. Most of the effectiveness of SRL interventions for improving student achievement is thus not due to improvements in SRL, but due to other factors, contrary to common assumption. Furthermore, we provided insight into the factors that moderate each of the three studied relations in higher education. We have thereby shown that SRL interventions may have different effects on students' achievement and students' engagement in SRL activities. By combining mediation analysis with three separate meta-analyses, this systematic review provides a thorough and complete synthesis of current SRL research in higher education, while opening new avenues for future exploration.

Acknowledgements

This work is financed via a grant by the Dutch National Initiative for Education Research (NRO)/The Netherlands Organization for Scientific Research (NWO) and the Dutch Ministry of Education, Culture and Science under grant nr. 405-15-705 (SOONER/ http://sooner.nu). Suzanne Jak was supported by a grant from The Netherlands Organization for Scientific Research: NWO-VENI-451-16-001.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.edurev.2019.100292.

References

References marked with an asterisk indicate studies included in the meta-analyses. A complete reference list of studies included in the meta-analyses is available as supplementary material (online only).

Artino, A. R. (2008). Motivational beliefs and perceptions of instructional quality: Predicting satisfaction with online training: Predicting satisfaction with online training. *Journal of Computer Assisted Learning, 24*(3), 260–270. https://doi.org/10.1111/j.1365-2729.2007.00258.x.

Artino, A. R., & Stephens, J. M. (2006). Learning online: Motivation to self-regulate? *Academic Exchange Quarterly, 10*, 176–182.

*Azevedo, R., & Cromley, J. G. (2004). Does training on self-regulated learning facilitate students' learning with hypermedia? *Journal of Educational Psychology, 96*, 523–535. https://doi.org/10.1037/0022-0663.96.3.523.

*Azevedo, R., Cromley, J. G., & Seibert, D. (2004). Does adaptive scaffolding facilitate students' ability to regulate their learning with hypermedia? *Contemporary Educational Psychology, 29*, 344–370. https://doi.org/10.1016/j.cedpsych.2003.09.002.

Banks, G. C., Kepes, S., & Banks, K. P. (2012). Publication bias: The antagonist of meta-analytic reviews and effective policymaking. *Educational Evaluation and Policy Analysis, 34*, 259–277. https://doi.org/10.3102/0162373712446144.

*Bannert, M., Hildebrand, M., & Mengelkamp, C. (2009). Effects of a metacognitive support device in learning environments. *Computers in Human Behavior, 25*, 829–835. https://doi.org/10.1016/j.chb.2008.07.002.
*Bannert, M., & Reimann, P. (2012). Supporting self-regulated hypermedia learning through prompts. *Instructional Science, 40*, 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. https://doi.org/10.1016/j.chb.2015.05.038.
Beishuizen, J., & Steffens, K. (2011). A conceptual framework for research on self-regulated learning. In R. Carneiro, P. Lefrere, K. Steffens, & J. Underwood (Eds.). *Self-regulated learning in technology enhanced learning environments* (pp. 3–19). Rotterdam, The Netherlands: Sense Publishers.
*Belski, R., & Belski, I. (2014). Cultivating student skills in self-regulated learning through evaluation of task complexity. *Teaching in Higher Education, 19*, 459–469. https://doi.org/10.1080/13562517.2014.880685.
*Berthold, K., Nückles, M., & Renkl, A. (2007). Do learning protocols support learning strategies and outcomes? The role of cognitive and metacognitive prompts. *Learning and Instruction, 17*(5), 564–577. https://doi.org/10.1016/j.learninstruc.2007.09.007.
Bertrams, A., & Schmeichel, B. J. (2014). Improving self-control by practicing logical reasoning. *Self and Identity, 13*, 419–431. https://doi.org/10.1080/15298868.2013.836562.
Boekaerts, M. (1992). The adaptable learning process: Initiating and maintaining behavioural change. *Applied Psychology: International Review, 41*, 377–397.
Boer, H. de, Donker-Bergstra, A. S., Kostons, D. D. N. M., & Korpershoek, H. (2013). *Effective strategies for self-regulated learning: A meta-analysis.* Groningen: GION/RUG.
*van den Boom, G., Paas, F., & van Merriënboer, J. J. G. (2007). Effects of elicited reflections combined with tutor or peer feedback on self-regulated learning and learning outcomes. *Learning and Instruction, 17*, 532–548. https://doi.org/10.1016/j.learninstruc.2007.09.003.
Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (Eds.). (2009). *Introduction to meta-analysis*. Chichester, UK: Wiley.
Borenstein, M., Hedges, L. V., Higgins, J. P. T., & Rothstein, H. R. (2014). *Comprehensive meta analysis version 3.* Englewood, NJ, USA: Biostat.
Broadbent, J., & Poon, W. L. (2015). Self-regulated learning strategies & academic achievement in online higher education learning environments: A systematic review. *The Internet and Higher Education, 27*, 1–13. https://doi.org/10.1016/j.iheduc.2015.04.007.
de Bruijn-Smolders, M., Timmers, C. F., Gawke, J. C. L., Schoonman, W., & Born, M. P. (2016). Effective self-regulatory processes in higher education: Research findings and future directions. A systematic review. *Studies in Higher Education, 41*, 139–158. https://doi.org/10.1080/03075079.2014.915302.
Burke, L. A., & Hutchins, H. M. (2007). Training transfer: An integrative literature review. *Human Resource Development Review, 6*, 263–296. https://doi.org/10.1177/1534484307303035.
Cheung, M. W.-L. (2014). Fixed- and random-effects meta-analytic structural equation modeling: Examples and analyses in R. *Behavior Research Methods, 46*, 29–40. https://doi.org/10.3758/s13428-013-0361-y.
Cheung, M. W.-L. (2015). metaSEM: an R package for meta-analysis using structural equation modeling. *Frontiers in Psychology, 5*https://doi.org/10.3389/fpsyg.2014.01521.
Clarebout, G., Horz, H., Schnotz, W., & Elen, J. (2010). The relation between self-regulation and the embedding of support in learning environments. *Educational Technology Research & Development, 58*, 573–587. https://doi.org/10.1007/s11423-009-9147-4.
*Cleary, T. J., Callan, G. L., Malatesta, J., & Adams, T. (2015). Examining the level of convergence among self-regulated learning microanalytic processes, achievement, and a self-report questionnaire. *Journal of Psychoeducational Assessment, 33*, 439–450. https://doi.org/10.1177/0734282915594739.
Cohen, J. (1988). *Statistical power analysis for the behavioral sciences.* New Jersey, NJ: Lawrence Erlbaum Associates.
Credé, M., & Phillips, L. A. (2011). A meta-analytic review of the motivated strategies for learning questionnaire. *Learning and Individual Differences, 21*(4), 337–346. https://doi.org/10.1016/j.lindif.2011.03.002.
*Delen, E., Liew, J., & Willson, V. (2014). Effects of interactivity and instructional scaffolding on learning: Self-regulation in online video-based environments. *Computers & Education, 78*, 312–320. https://doi.org/10.1016/j.compedu.2014.06.018.
Dent, A. L., & Koenka, A. C. (2016). The relation between self-regulated learning and academic achievement across childhood and adolescence: A meta-analysis. *Educational Psychology Review, 28*, 425–474. https://doi.org/10.1007/s10648-015-9320-8.
Devolder, 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: Scaffolding self-regulated learning with CBLES. *Journal of Computer Assisted Learning, 28*, 557–573. https://doi.org/10.1111/j.1365-2729.2011.00476.x.
Dignath, C., Buettner, G., & Langfeldt, H.-P. (2008). How can primary school students learn self-regulated learning strategies most effectively? *Educational Research Review, 3*, 101–129. https://doi.org/10.1016/j.edurev.2008.02.003.
Dignath, C., & Büttner, G. (2008). Components of fostering self-regulated learning among students. A meta-analysis on intervention studies at primary and secondary school level. *Metacognition and Learning, 3*, 231–264. https://doi.org/10.1007/s11409-008-9029-x.
*Dörrenbächer, L., & Perels, F. (2016). More is more? Evaluation of interventions to foster self-regulated learning in college. *International Journal of Educational Research, 78*, 50–65. https://doi.org/10.1016/j.ijer.2016.05.010.
Dunlosky, J., Rawson, K. A., Marsh, E. J., Nathan, M. J., & Willingham, D. T. (2013). Improving students' learning with effective learning techniques: Promising directions from cognitive and educational psychology. *Psychological Science in the Public Interest, 14*(1), 4–58. https://doi.org/10.1177/1529100612453266.
Efklides, A. (2011). Interactions of metacognition with motivation and affect in self-regulated learning: The MASRL model. *Educational Psychologist, 46*, 6–25. https://doi.org/10.1080/00461520.2011.538645.
Feinstein, A. R., & Cicchetti, D. V. (1990). High agreement but low kappa: I. The problems of two paradoxes. *Journal of Clinical Epidemiology, 43*, 543–549. https://doi.org/10.1016/0895-4356(90)90158-L.
Fisher, Z., Tipton, E., & Zhipeng, H. (2017). *robumeta: Robust variance meta-regression (version 2.0)*. Retrieved from https://CRAN.R-project.org/package=robumeta.
*Garavalia, L. S., & Gredler, M. E. (2002). An exploratory study of academic goal setting, achievement calibration, and self-regulated learning. *Journal of Instructional Psychology, 29*, 221–230.
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.
*Greene, J. A., Hutchison, L. A., Costa, L.-J., & Crompton, H. (2012). Investigating how college students' task definitions and plans relate to self-regulated learning processing and understanding of a complex science topic. *Contemporary Educational Psychology, 37*, 307–320. https://doi.org/10.1016/j.cedpsych.2012.02.002.
Hadwin, A. F., Nesbit, J. C., Jamieson-Noel, D., Code, J., & Winne, P. H. (2007). Examining trace data to explore self-regulated learning. *Metacognition and Learning, 2*, 107–124. https://doi.org/10.1007/s11409-007-9016-7.
Hattie, J., Biggs, J., & Purdie, N. (1996). Effects of learning skills interventions on student learning: A meta-analysis. *Review of Educational Research, 66*, 99–136. https://doi.org/10.3102/00346543066002099.
Hedges, L. V., Tipton, E., & Johnson, M. C. (2010). Robust variance estimation in meta-regression with dependent effect size estimates. *Research Synthesis Methods, 1*, 39–65. https://doi.org/10.1002/jrsm.5.
Higgins, J. P. T., & Thompson, S. G. (2002). Quantifying heterogeneity in a meta-analysis. *Statistics in Medicine, 21*, 1539–1558. https://doi.org/10.1002/sim.1186.
Hox, J. J. (2010). The multilevel approach to meta-analysis. *Multilevel analysis: Techniques and applications* (pp. 205–232). (2nd ed.). New York, NY: Routledge.
Jak, S. (2015). *Meta-analytic structural equation modelling.* Cham: Springer International Publishing. https://doi.org/10.1007/978-3-319-27174-3.
*Lusk, A. (2016). Metacognitive strategies in the introduction to political science classroom. *Journal of Political Science Education, 12*, 141–150. https://doi.org/10.1080/15512169.2015.1060888.
*Masui, C., & De Corte, E. (2005). Learning to reflect and to attribute constructively as basic components of self-regulated learning. *British Journal of Educational Psychology, 75*, 351–372. https://doi.org/10.1348/000709905X25030.
*McKenzie, K., Gow, K., & Schweitzer, R. (2004). Exploring first-year academic achievement through structural equation modelling. *Higher Education Research and Development, 23*, 95–112. https://doi.org/10.1080/0729436032000168513.
*Moos, D. C., & Bonde, C. (2016). Flipping the classroom: Embedding self-regulated learning prompts in videos. *Technology, Knowledge and Learning, 21*, 225–242.

https://doi.org/10.1007/s10758-015-9269-1.
Nelson, T. O., & Narens, L. (1990). Metamemory: A theoretical framework and new findings. *Psychology of learning and motivation: Vol.26*, (pp. 125–173). Academic Press.
*Nietfeld, J. L., Cao, L., & Osborne, J. W. (2006). The effect of distributed monitoring exercises and feedback on performance, monitoring accuracy, and self-efficacy. *Metacognition and Learning,* 159–179. https://doi.org/10.1007/s10409-006-9595-6.
*Ning, H. K., & Downing, K. (2012). Influence of student learning experience on academic performance: The mediator and moderator effects of self-regulation and motivation. *British Educational Research Journal, 38*, 219–237. https://doi.org/10.1080/01411926.2010.538468.
*Nückles, M., Hübner, S., & Renkl, A. (2009). Enhancing self-regulated learning by writing learning protocols. *Learning and Instruction, 19*, 259–271. https://doi.org/10.1016/j.learninstruc.2008.05.002.
Orwin, R. G. (1983). A fail-safe N for effect size in meta-analysis. *Journal of Educational and Behavioral Statistics, 8*, 157–159. https://doi.org/10.3102/10769986008002157.
Panadero, E. (2017). A review of self-regulated learning: Six models and four directions for research. *Frontiers in Psychology, 8*https://doi.org/10.3389/fpsyg.2017.00422.
Panadero, E., Klug, J., & Järvelä, S. (2016). Third wave of measurement in the self-regulated learning field: When measurement and intervention come hand in hand. *Scandinavian Journal of Educational Research, 60*(6), 723–735. https://doi.org/10.1080/00313831.2015.1066436.
Paris, S. G., & Paris, A. H. (2001). Classroom Applications of research on self-regulated learning. *Educational Psychologist, 36*(2), 89–101. https://doi.org/10.1207/S15326985EP3602_4.
Perels, F., Gürtler, T., & Schmitz, B. (2005). Training of self-regulatory and problem-solving competence. *Learning and Instruction, 15*, 123–139. https://doi.org/10.1016/j.learninstruc.2005.04.010.
Pieger, E., & Bannert, M. (2018). Differential effects of students' self-directed metacognitive prompts. *Computers in Human Behavior, 86*, 165–173. https://doi.org/10.1016/j.chb.2018.04.022.
Pintrich, P. R. (1999). The role of motivation in promoting and sustaining self-regulated learning. *International Journal of Educational Research, 31*, 459–470. https://doi.org/10.1016/S0883-0355(99)00015-4.
Pintrich, P. R. (2000). The role of goal orientation in self-regulated learning. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.). *Handbook of self-regulation* (pp. 451–502). San Diego, CA, US: Academic Press.
Pintrich, P. R. (2004). A conceptual framework for assessing motivation and self-regulated learning in college students. *Educational Psychology Review, 16*, 385–407. https://doi.org/10.1007/s10648-004-0006-x.
Puustinen, M., & Pulkkinen, L. (2001). Models of self-regulated learning: A review. *Scandinavian Journal of Educational Research, 45*, 269–286. https://doi.org/10.1080/00313830120074206.
R Core Team (2016). *R: A language and environment for statistical computing.* Vienna, Austria: R Foundation for Statistical Computing. Retrieved from https://www.R-project.org/.
*Schmidt, A. M., & Ford, J. K. (2003). Learning within a learner control training environment: The interactive effects of goal orientation and metacognitive instruction on learning outcomes. *Personnel Psychology, 56*, 405–429. https://doi.org/10.1111/j.1744-6570.2003.tb00156.x.
Schnotz, W. (2010). Reanalyzing the expertise reversal effect. *Instructional Science, 38*(3), 315–323. https://doi.org/10.1007/s11251-009-9104-y.
Schunk, D. H. (2005). Self-regulated learning: The educational legacy of Paul R. Pintrich. *Educational Psychologist, 40*(2), 85–94. https://doi.org/10.1207/s15326985ep4002_3.
*Sitzmann, T., Bell, B. S., Kraiger, K., & Kanar, A. M. (2009). A multilevel analysis of the effect of prompting self-regulation in technology-delivered instruction. *Personnel Psychology, 62*, 697–734. https://doi.org/10.1111/j.1744-6570.2009.01155.x.
Sitzmann, T., & Ely, K. (2011). A meta-analysis of self-regulated learning in work-related training and educational attainment: What we know and where we need to go. *Psychological Bulletin, 137*, 421–442. https://doi.org/10.1037/a0022777.
*Stark, R., & Krause, U.-M. (2009). Effects of reflection prompts on learning outcomes and learning behaviour in statistics education. *Learning Environments Research, 12*, 209–223. https://doi.org/10.1007/s10984-009-9063-x.
Van Laer, S., & Elen, J. (2017). In search of attributes that support self-regulation in blended learning environments. *Education and Information Technologies, 22*, 1395–1454. https://doi.org/10.1007/s10639-016-9505-x.
Van Merriënboer, J. J. G., & Sweller, J. (2010). Cognitive load theory in health professional education: Design principles and strategies: Cognitive load theory. *Medical Education, 44*, 85–93. https://doi.org/10.1111/j.1365-2923.2009.03498.x.
Veenman, M. V. J. (2005). The assessment of metacognitive skills: What can be learned from multi-method design? In C. Artelt, & B. Moschner (Eds.). *Lernstrategien und Metakognition: Implikationen für Forschung und Praxis* (pp. 75–97). Berlin: Waxmann.
Veenman, M. V. J. (2016). Learning to self-monitor and self-regulate. In R. E. Mayer, & P. A. Alexander (Eds.). *Handbook of research on learning and instruction* (pp. 233–257). New York, NY: Routledge.
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*(1), 3–14. https://doi.org/10.1007/s11409-006-6893-0.
Wang, C.-H., Shannon, D. M., & Ross, M. E. (2013). Students' characteristics, self-regulated learning, technology self-efficacy, and course outcomes in online learning. *Distance Education, 34*, 302–323. https://doi.org/10.1080/01587919.2013.835779.
Winne, P. H. (2010). Improving measurements of self-regulated learning. *Educational Psychologist, 45*(4), 267–276. https://doi.org/10.1080/00461520.2010.517150.
Winne, P. H., & Hadwin, A. F. (1998). Studying as self-regulated learning. In D. J. Hacker, J. Dunlosky, & A. Graesser (Eds.). *Metacognition in educational theory and practice* (pp. 277–304). Mahwah, NJ, US: Lawrence Erlbaum Associates.
Winne, P. H., Jamieson-Noel, D., & Muis, K. R. (2002). Methodological issues and advances in researching tactics, strategies, and self-regulated learning. In (1. ed). P. R. Pintrich, & M. L. Maehr (Vol. Eds.), *New directions in measures and methods: Vol.12*, (pp. 121–155). Amsterdam: JAI, An Imprint of Elsevier Science.
Winne, P. H., & Perry, N. E. (2000). Measuring self-regulated learning. In M. Boekaerts, P. R. Pintrich, & M. Zeidner (Eds.). *Handbook of self-regulation* (pp. 531–566). Elsevier.
*Wischgoll, A. (2016). Combined training of one cognitive and one metacognitive strategy improves academic writing skills. *Frontiers in Psychology, 7*, 1–13. https://doi.org/10.3389/fpsyg.2016.00187.
Young, J. Q., Van Merriënboer, J. J. G., Durning, S., & Ten Cate, O. (2014). Cognitive load theory: Implications for medical education: AMEE guide No. 86. *Medical Teacher, 36*(5), 371–384. https://doi.org/10.3109/0142159X.2014.889290.
Zimmerman, B. J. (1986). Becoming a self-regulated learner: Which are the key subprocesses? *Contemporary Educational Psychology, 11*(4), 307–313. https://doi.org/10.1016/0361-476X(86)90027-5.
Zimmerman, B. J. (1990). Self-regulated learning and academic achievement: An overview. *Educational Psychologist, 25*, 3–17. https://doi.org/10.1207/s15326985ep2501_2.
Zimmerman, B. J. (2002). Becoming a self-regulated learner: An overview. *Theory into Practice, 41*, 64–70. https://doi.org/10.1207/s15430421tip4102_2.
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.