Does training in how to regulate one's learning affect how students report self-regulated learning in diary tasks?

P. Costa Ferreira · A. M. Veiga Simão · A. Lopes da Silva

Abstract The processes and perceptions of students' self-regulated learning are not easily measured. Thus, research has presented and suggested numerous ways in which these processes and perceptions of self-regulated learning can be investigated and assessed. Accordingly, this study aims to assess whether training in how to regulate one's learning is related to students' growth patterns regarding their reported self-regulated learning activity over time. This study also investigates whether this type of training has an impact on students' reflective ability and academic performance. To reach these goals, we examined whether students' use of a diary task - developed by interviewing primary school students ($n=43$) and validated with exploratory ($n=78$) and confirmatory ($n=83$) factor analysis - would capture change in students' reported self-regulated learning activity and reflective ability during training in how to regulate one's learning (students: $n=100$; diary task entries: $n=1,000$). Students' academic performance was assessed with an oral and vocabulary task. Results from multilevel linear modeling revealed different growth rates of reported self-regulated learning activity over time between students who experienced training and students who did not. Furthermore, pre and posttest results revealed that the students who experienced the training reported their reflections more autonomously and specifically in their diary task and had better academic performance than students who did not. These results demonstrate how the diary task captured change in students' perceptions, validating it as a monitoring tool. Lastly, implications for practitioners are discussed and suggestions for future studies are proposed.

Keywords Self-regulated learning · Assessment · Learning diary task · Multilevel linear models

---

P. C. Ferreira · A. M. V. Simão · A. L. da Silva
Faculty of Psychology, University of Lisbon, Alameda da Universidade, 1649-013 Lisbon, Portugal

P. C. Ferreira (✉)
The Portuguese Foundation for Science and Technology (FCT), Lisbon, Portugal
e-mail: paula.costa.ferreira@gmail.com

Introduction

The regulation of learning is a fundamental requirement for the successful attainment of knowledge and skills in academic contexts and moreover, in life-long learning. It implies reflection, affect and action that are developed by learners in order to reach academic goals. Thus, these goals are pursued by learners proactively through learning which is self-oriented (Zimmerman 2000). In the self-regulation of learning, learners become actively involved in self-awareness, self-motivation, and behavioral competence in order to develop the capacity to construct and employ knowledge appropriately (Wolters et al. 2003). Ultimately, meaningful learning depends greatly on the deliberate construction of knowledge and use of effective learning strategies in any context (McNamara 2011). This study investigated whether training in how to regulate one's learning could somehow influence students' growth patterns of their reported self-regulated learning activity over time, their ability to reflect about their own functioning in class, and their academic performance. We use the term regulation of learning, as opposed to self-regulated learning when referring to training because students do not work and regulate their learning alone (self-regulate) in a classroom environment, but in collaboration with others (i.e. co- and shared regulation) (Järvelä and Hadwin 2013; Määttä et al. 2012). Hence, in this study we use the term self-regulated learning when we refer to students' reports and reflections of their own functioning.

Zimmerman (2000) defined self-regulated learning as the degree with which learners metacognitively, motivationally, and behaviorally manage their own learning process. Particularly, learners are metacognitively aware and motivationally connected to how they regulate their learning by actively adapting strategies to execute specific learning tasks. Additionally, Zimmerman presented the process of regulating one's own learning in three cyclical self-regulatory phases. Specifically, the forethought phase, where learners set objectives and plan before a task, the performance phase, where learners monitor and control their performance while they execute the task, and the self-reflection phase, where learners react to their own outcomes once the learning process is completed. As Zimmerman stated, these phases may help clarify learners' repeated efforts to learn in terms of quantitative and qualitative differences (i.e., proactive vs. reactive self-regulators), such as when learning a new language (see Zimmerman 2013). Zimmerman and Martinez-Pons (1986) found several strategies for regulating learning that are associated with covert (i.e., goal setting and planning, organizing and transforming, seeking information, and rehearsing and memorizing), behavioral (i.e., keeping records and monitoring, reviewing records, and self-evaluation), and environmental (i.e., environmental structuring, seeking social assistance, self-consequences, and environmental structuring) forms of self-regulated learning. In the present study, students were taught these strategies in a sequence to help them develop skills that would allow them to plan (forethought phase), monitor (performance phase) and self-evaluate (reflection phase) their performance.

The literature regarding the regulation of learning in classroom contexts has suggested that research should focus on how the processes involved in the regulation of learning are taught (encompassing goal establishment, assessment of learner beliefs about learning and self-evaluation), as well as on the precision with which learners report and register their use of self-regulatory processes (e.g. Winne 2011; Zimmerman 2008). Bandura (2006) proposed how learners' behavior is conditioned by external aspects and by their beliefs about how they can perform in specific domains. Thus, Bandura posited how it is also imperative that learners learn to self-monitor and regulate their thoughts, emotions and actions. Accordingly, since some studies have raised questions about existing instruments and their ability to capture self-regulated learning processes and perceptions about those processes as they occur, research

should focus on process studies using different tools, such as diary tasks (e.g. Azevedo and Cromley 2004; Schmitz 2006; Schmitz et al. 2011).

Cleary (2011) suggested the use of microanalysis as a way to measure learners' regulatory beliefs and reactions as they engage in academic tasks. According to Cleary, self-regulated learning is viewed by most scholars as dynamic and fluid, rather than static. Hence, micro-analysis and the use of diary tasks for instance, are alternative approaches to standard scales, which help interpret the regulation of learning as a global construct. Microanalysis refers to a detailed form of measuring specific processes and behaviors as they occur "in real time across authentic contexts" (Cleary 2011, p. 330). Moreover, self-regulated learning microanalysis focuses primarily on the beliefs and processes involved in self-regulation which are often covert. These beliefs and processes are examined through the use of structured assessment tools that evaluate the cyclical phases of self-regulated learning "at strategic moments during a specific activity" (Cleary 2011, p. 331). Diaries, for instance, are able to capture changes that may occur in these processes and behaviors (Schmitz 2006). Cleary (2011) suggested important aspects of contemporary microanalytic assessment protocols, such as individual adminis-tration, the study of the three self-regulated learning phases proposed by Zimmerman (2000), connecting these phases to the before, during and after moments of an event, context-specific microanalytic questions, and lastly, recordings and coding of learners' responses.

In this study, we introduced a diary task that could provide detailed information about students' awareness of their self-regulated learning activity and that could capture change in students' reflections about learning. As Boekaerts (2002) and Hofer (2005) stated, teachers should have a realistic view of students' beliefs and perceptions regarding learning when planning learning activities. Moreover, in light of the rising learning demands of the 21st century, students must engage proactively and strategically in their academic path. We argue that students can begin working from early on in their academic career in order to contribute to their life-long learning process efficaciously. The development and adaptation of motivation, behavior and new competencies are increasingly important for students to flexibly and auton-omously manage their learning. Hence, we also presented an example of training in how to regulate one's learning, as a potential guide for teachers when designing meaningful learning environments through the adaptation of pedagogical practices to attend to their students' needs.

Thus, we first examined students' reported self-regulated learning activity (in the diary task) over time, and assessed whether training in how to regulate one's learning could be related to their growth patterns. By reported self-regulated learning activity, and in accordance with Zimmerman's model of self-regulated learning (2000) and Bandura's theory of Human Agency (2006), we mean that students are self-regulators and once they have adopted intentions and action plans, they build a course of action, motivate and regulate their execution (self-regulated learning activity), connecting thought and action. Therefore, students in this study reported their immediate self-reactions about their self-regulated learning activity in class.

Secondly, this study investigated whether the training in how to regulate one's learning had an impact on students' reflective ability (i.e. reported intentions to learn, reported anticipations of learning outcomes and reported self-examination). Specifically by reflective ability we mean, and again considering Bandura (2006) and Zimmerman's (2000) work, that students are planners with a deliberate capacity to make choices and set goals, they are forethinkers who anticipate outcomes of prospective actions, and self-examiners who have the metacognitive capability to examine their own functioning by reflecting on their own efficacy, appropriate-ness of thoughts and actions, as well as on the value of their own efforts.

Thirdly, this study investigated whether the training in how to regulate one's learning had an impact on students' academic performance. We also discuss in this paper how this training and using a diary task that follows contemporary microanalytic assessment protocol guidelines

may be an important approach to measure students' reported self-regulated learning activity and reflective ability. Specifically, we proposed to answer the following questions:

Are there differences in growth rates of reported self-regulated learning activity over time between students who experience training and students who do not?
Do students who experience training report their reflections more autonomously and specifically in their diary task than students who do not?
Do students who experience training have better academic performance than students who do not?

Children's self-monitoring with diary tasks in domain-specific contexts

Self-monitoring is the ability learners have to pay attention to and closely observe their own learning behavior at a meta-level in order to verify and/or adjust their management of learning during present and future efforts to learn (Zimmerman 1989). Winne (2011) posited that because metacognitive monitoring guides learners' choices about learning, it is vital to understand their perceptions of the learning context. What's more, and in line with Klug et al. (2011), monitoring is a high-order operation of self-regulated learning in the sense that it provides learners with feedback regarding purposeful behavior and learning outcomes. Self-monitoring can be fostered in children and other age groups through self-recording with the use of standardized learning diary tasks, which in turn, trigger a positive reactivity effect, leading to deep reflection about their own learning behavior and learning process and thus, directing them towards better academic performance (Schmitz and Wiese 2006). Current research has proposed that children aged 8 to 10 have the capacity to regulate their learning, while using metacognitive processes consistently and maturely (e.g. Bares 2011). What's more, the literature has also suggested that children in this age group are also capable of making accurate and differential judgments regarding their self-regulation competencies when the evaluation method is appropriate for their age (Rizzo et al. 2010).

Recent studies have revealed how learning diary tasks stimulate learners to reflect on, inquire about and gain awareness of their learning experiences (Alterio 2004; Gbabremani-Gbajar and Mirbosseini 2005; Simard 2004). Additional evidence has shown that the act of registering experiences in a learning diary task can capture occurrences and reveal feelings and thoughts during the process (Boud 2001). What's more, empirical findings have demonstrated that learning diary tasks are effective cognitive tools and/or learning methods that provide learners with important insights into how they perceive their learning experiences and can ultimately help them reflect on new acquired information (Hiemstra 2001; Kaur 2003). Further studies have sought to understand how learning diary tasks can be effective with children (4th grade) along with interventions in how to regulate learning, providing thus a better development of self-regulated learning processes and improved academic performance (Otto 2007; Stoeger and Ziegler 2010).

Some authors have highlighted how interventions in regulating learning as a process can be appropriate for different age groups (Klug et al. 2011) and for elementary school children in particular (Glaser and Brunstein 2007). Schmitz and Perels (2011) for instance, investigated the effect of an intervention in self-regulated learning on 8th grade students who filled in diary tasks for a period of 49 days during math homework activities. Through trend analysis, the authors found that there was a highly significant positive linear trend in favor of the group that received the training. Also, through pre-posttest measures, the authors found an increased self-regulation in the students that experienced guidance in how to regulate learning.

In view of these studies, and in an attempt to answer our questions about the effects of training in how to regulate one's learning on how students report their self-regulated learning activity in their diary task, we present our first hypothesis. Hence, we hypothesize that:

H1: *There are different growth rates of reported self-regulated learning activity over time between students who experience training and students who do not.*

Stoeger and Ziegler (2008) conducted a 5-week intervention study where 4th grade teachers gave their students training in how to regulate learning in math classes. With the data from the diary tasks and other pre and posttest instruments such as questionnaires, the authors found that students who experienced the training revealed a greater increase in effort, task interest, learning-goal orientation, and perceptions of self-efficacy, than the students in the control-group. In a later study, the same authors (Stoeger and Ziegler 2011) found similar results with the implementation of a training module on how to regulate one's learning on 4th grade students studying math for 6 weeks. Through hierarchical linear modeling analyses, the authors found that students who experienced training in self-regulated learning with the use of learning diary tasks, displayed a greater level of learning-goal orientation, task interest, effort and perceptions of self-efficacy, than learners who did not. Hence, training in how to regulate one's learning was successful and was recommended for learners as young as those in elementary school.

Perels et al. (2007) analyzed self-regulation from a process perspective and studied the impact an intervention program had on the development of self-regulation with data gathered from 8th grade students' diary tasks. The authors found an impact of the training regarding planning, resource use, learning environment structuring and distraction management through time series analysis, as well as significant trends for starting to learn at a planned time, self-monitoring, and self-reflection with trend analyses.

We believe that through the type of impact that Perels and her colleagues (2007) found, students' intentions to learn, expectations of learning outcomes and self-examinations/reflections about past experiences may also be affected. Thus, we expect that training in how to regulate one's learning affects how students report their reflections as strategic and autonomous agents. Accordingly, we assume that:

H2: *Students who experience training report their reflections more autonomously and specifically in their diary task than students who do not.*

Still regarding self-regulation aspects, but also focusing on academic performance, Perels et al. (2009) worked with 6th grade students using diary tasks as well, and investigated whether training in how to regulate one's learning could influence students' capacity to self-regulate and to achieve better academic performance in math. By means of analyses of variance with time as a repeated measurement factor and analyses of covariances, the authors found that there were significant interactions between time and group for most of the self-regulated learning variables in favor of the group that had training. Essentially, this study demonstrated that training in how to regulate one's learning could support self-regulation competencies and math achievement, thus revealing higher scores for the experimental group.

Glogger et al. (2012) analyzed whether the quantity and quality of learning strategies measured by diary tasks could predict learning outcomes and whether different effective combinations of strategies could be identified. With the responses of 9th grade students in math and biology classes for a period of 6 weeks, the authors found through hierarchical linear modeling that the quality and quantity of cognitive strategies predicted learning outcomes. The authors also found through cluster analysis that learners who combined cognitive and meta cognitive strategies were more successful. Overall, the authors concluded that using diary tasks was a useful way of assessing self-regulated learning strategies.

Otto (2007) worked with 4th grade students and investigated a 5-week intervention program which aimed to promote the regulation of learning in mathematical problem solving. Students were divided into five groups, namely, a simultaneous student, parent and teacher training group, a simultaneous student and instructor training group, a simultaneous parent and teacher training group, a teacher training group and a non-training control group. Through longitudinal (pretest, posttest, stability measurement) and procedural data (trend analysis of standardized learning diary tasks), the author found significant positive effects regarding self-regulated learning strategies and performance in math, mostly in regards to the group of students who had direct self-regulated learning training and used the diary tasks.

In the present study, students participated with diary tasks and had training on how to regulate one's learning in English as a foreign language (EFL) class. Learning a foreign language implies learning a language other than the learner's first language, the language of instruction or a second language, which is the language of instruction for learners with a minority background (European Commission 2013). Oxford (2003) indicates how learning a foreign language efficiently is in part determined by learners' learning styles and strategies. Furthermore, language learning strategies that are chosen intentionally, function as a means for conscious regulation of learning. Hence, being aware of how to learn a language more efficiently by practicing goal-setting, self-monitoring and self-evaluation can lead to autonomous language learning.

In a theoretical and empirical review of studies focusing on metacognitive knowledge and language learning, Rahimi and Katal (2012) discussed how learners who are aware and act consciously to understand what they are doing by using different strategies to plan, monitor and evaluate their learning, are usually the most successful learners in foreign language learning. Some authors (Kirsh 2012; Larkin 2010) have argued that providing young learners with opportunities for reflection while learning a foreign language, such as English, enables them to enhance their strategy use, develop self-regulation, autonomy and proficiency in using the language. Specifically, Larkin (2010) provided evidence that with time and reflection, learners are capable of developing metacognitive and metalinguistic awareness, allowing them to improve their use of the language. Hence, providing learners with tools that allow them to reflect, is imperative for effective language learning. Some of the literature on learning EFL has indicated that using diary tasks can be beneficial with young learners because they can be a good tool for reflection and awareness of learning interests and difficulties (Kir 2012). Nonetheless, Rahimi and Katal (2012) suggested how more research is needed to understand how learners can become metacognitively aware and efficient regulators of their own language learning in EFL settings. Accordingly, Greene et al. (2010) claimed that research needs to study the regulation of learning processes in different domain-areas other than math and sciences (areas which have been focused on in the vast literature on the regulation of learning). Thus, we also decided to focus on the academic performance of students in an EFL context (i.e., oral practice and vocabulary in EFL, as described in the method section) and hypothesized that:

H3: *Students who experience training will have better academic performance in EFL than students who do not.*

Method

Study design

A quasi-experimental control-group design with repeated measurements (with pre and posttest) was used with process data gathered in a real life context (Klug et al. 2011).
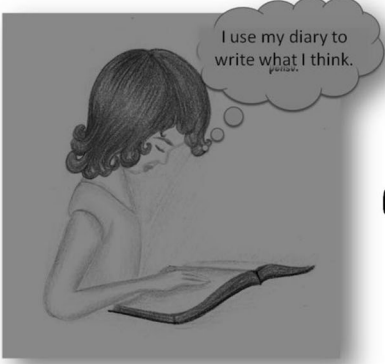
Participants

The convenience sample used in this study consisted of a total of 204 students in EFL classes. All participating students were of middle class families (94 % Portuguese, 4 % Brazilian and 2 % Angolan), all spoke Portuguese fluently and all were in the fourth grade of different schools in the Lisbon area. Also, all were in the same level of EFL (A1 level – basic user) according to teachers' indications, which were in line with the Common European Framework of Reference for Languages (Council of Europe 2011). While 43 ($M=9.23$ years, $SD=.84$, 54 % boys) students participated in the preliminary development of the quantitative items of the diary task, 78 ($M=9.3$ years, $SD=.51$, 55 % boys) participated in the first pilot study (exploratory factor analysis) and 83 ($M=9.4$ years, $SD=.62$, 54 % girls) participated in the second pilot study (confirmatory factor analysis). A total of 100 students ($M=9.2$ years, $SD=.42$, 53 % boys), with their respective teachers participated in the main study presented here, allowing us to gather 1,000 diary task entries for data analysis with multilevel linear modeling. The experimental group was composed of 40 students, while the control group was composed of 60 students. This distribution was based on the willingness of the groups' respective teachers to participate in the training sessions.

Instruments

Diary task The Diary of Guided Self-regulated learning (DOGS-RL) is based on the literature mentioned (e.g. Bandura 2006; Schmitz and Wiese 2006; Schunk and Zimmerman 1998; Zimmerman and Martinez-Pons 1986) and consists of a total of three open-ended questions, three dichotomous questions and 12 quantitative items that should be completed in class individually. The open-ended qualitative items of the diary task ($\alpha=.80$) (i.e., 'Today in my class I want to learn…'; 'In my class today I will be able to…'; 'Today in my class I learned…') are based on recommendations in the literature (Bandura and Locke 2003; Cleary 2011; Zimmerman 1989), and specifically in the areas of Agency of Intentionality, Forethought and Self-reflectiveness (Bandura 2006), which fit in with the various subprocesses of self-regulated learning (Cleary 2011; Zimmerman 2008). The scores for these questions were developed according to guidelines in the literature (Kember 2004; Kember et al. 1999; Kember et al. 2008; Wong et al. 1995) and students' responses. Specifically, the first part of the diary task (see Fig. 1) begins with two open-ended questions that are meant to be responded before efforts to learn in class. The first question asks about the students' intentions to learn ("Today in my class I want to learn…" coded on a scale of 9 points: from 1=irrelevant comments to 9=specific and autonomous goal regarding strategic action) and the second asks about the students' anticipations of their learning outcomes/performance ("Today in my class I will be able to…" coded on a scale of seven points: from 1=irrelevant comments to 7=specific anticipation regarding strategic action).

The second part of the diary task is meant to be responded to immediately after students engage in learning activities in class (see Fig. 2). This second part includes three blocks of questions, the first about forethought, the second about performance and the third about self-reflection (three cyclical phases of self-regulated learning), as well as a last open-ended question. The block of questions about forethought begins with one dichotomous question ("Did I plan my work in class today?") which is to be responded on the basis of yes or no. If the response is yes, the students move onto the next question about planning ("How?") to help them focus on what they actually did in context. Accordingly, they then respond to the four quantitative items about forethought ("I liked to plan my work"; "I found it difficult to plan my work"; "I made an effort to plan my work"; and "I was able to plan my work") from 1 (not at

Fig. 1 First part of the diary task with the first and second open-ended questions

all) to 4 (a lot). The remaining two blocks of questions about performance ("Did I correct my work as I did it in class today?") and self-reflection ("Did I evaluate my work in class today?") present the same structure. Lastly, the diary presents one last open-ended question ("Today in my class I learned…" coded on a scale of 7 points: from 1=irrelevant comments to 7=specific cognitive and/or affective self-reflection regarding strategic action), where students report their self-examinations.

The full structure of this diary task asks students about all of the phases of the self-regulated learning cycle, including forethought, performance, and self-reflection in accordance with contemporary microanalytic assessment protocols as suggested in the literature (Cleary 2011).

Other measures The purpose of using an English oral task and vocabulary task about sports was to measure students' academic performance, since oral and reading skills are the most predominant in EFL classes in fourth grade in Portugal. An oral task (α=.84) was administered to students, where they had to describe in English what they saw in 10 different pictures about sports. Students were rated from 1 to 5 on the basis of whether they said no words (1), non-related words (2), related words (3), related sentences with mistakes (4), and related sentences with no mistakes (5). Also, students answered a 33 item English vocabulary test online about sports (α=.85), where they had to identify objects in the pictures and were scored dichotomously on whether they got the answer correct or incorrect. Both of these measures

Fig. 2 Second part of the diary task with the three blocks of quantitative items for forethought, performance and self-reflection, as well as the third open-ended question

were developed according to the content in the national curriculum of EFL for primary education and with the help of two primary school teachers. We checked for face and content validity of these measures by interviewing two other teachers regarding specific content-related questions (i.e., "What does the task assess?"; "What"s the objective of the task?" and "Do you think it is suitable for these students… if not, make suggestions.").

Procedures

Developing and testing the diary task In order to construct the quantitative items of the diary task, we interviewed 43 fourth-grade students from a different academic community based on previous theory and research (Zimmerman and Martinez-Pons 1986; Zimmerman 2008). The Self-regulated Learning Interview Schedule (Zimmerman and Martinez-Pons 1986) was translated and adapted (with a forward and back-translation of two bilingual researchers in the area of self-regulated learning and with pre-testing and cognitive interviewing of 3 volunteer students) to the students' academic conditions in fourth grade according to APA guidelines (Gudmundson 2009). Because we wanted to implement the diary task only inside the classroom, we only focused on two specific learning contexts, namely classroom learning (question 1 of the SRLIS: "Assume your teacher is discussing a topic with your class, such as World War

II, and he or she says that you will be tested on the topic. Do you have a particular method to help you learn and remember what was discussed in class?") and self-evaluation (question 8 of the SRLIS: "When taking tests in English, science or history, do you have a particular method for making sure your answers are correct before handing in the paper?"). Students' responses were first analyzed according to Zimmerman and Martinez-Pons' (1986) recommendations. Various strategies in classroom learning were identified as occurring frequently, namely: rehearsing and memorizing (e.g. "I am able to learn the material by practicing it a lot of times and correcting myself as I read it"), seeking information (e.g. "First I can do what is more difficult and search for it in my notebook."), reviewing records (e.g. "I like to reread my notes in order to learn and be prepared"), keeping records and monitoring (e.g. "I make an effort to take down notes of the difficult things in general and study them") and organizing and transforming (e.g. "First I like to plan my work by making an outline of the material so that I can learn it"). As for the context of self-evaluation, students mainly referred to self-evaluation (e.g. "I like to check if my text is good and I can correct what is wrong"; "It is difficult, but if I have time, I can correct my work"), and rehearsing and memorizing (e.g. "I can remember the material and try to concentrate so I can see if I am correct").

With the help of two primary school teachers, we performed content analysis of the students' responses in order to design the diary task. Common key-processes were identified in the students' discourse, such as, "liking to do something", "finding something difficult", "making an effort" and "being able to accomplish something". An initial 28 items were developed. After careful consideration with the two teachers in terms of item construction and difficulty of interpretation on the students' behalf, only 12 were considered for the diary task. First, students had to decide whether or not they had performed the action mentioned and answer no or yes. From there, if they answered yes, they would continue and answer not at all (1); a little (2); enough (3); or a lot (4) for each item pertaining to that action. The items were categorized into groups depending on the self-regulatory action (e.g. "I was able to prepare my work in class today.", "I found it difficult to correct my work as I did it in class today.", " I made an effort to evaluate my work in class today."). We wanted students to reflect on activities that they had just performed in class with the diary task, therefore, we used the simple past tense for all items (i.e., "Did I plan my work in class today?.").

We did a first pilot study (exploratory factor analysis) with 78 students to test the internal structure of the dairy task. For a confirmation of our results we did a second pilot study and asked 83 students to fill in the diary task. We tested our model with confirmatory factor analysis. To understand how students would react to the instrument in terms of motivation to respond during a prolonged period of time and difficulty in interpreting instructions, we had all students fill in the diary task during 10 lessons. This allowed us to understand the difficulties that students had in filling in the diary task and to perfect the instructions provided to students on how to fill in the instrument. Nonetheless, both in the exploratory factor analysis and in the confirmatory factor analysis, we opted to use students' responses from the last session.

The main study Once the diary task had been developed, we used it in our main study. We applied it in 12 EFL classes of 45 minutes twice a week over a period of 6 weeks on 100 participants. We opted for this number of weeks as in previous studies with interventions in how to regulate one's learning and diary task data (Otto 2007; Stoeger and Ziegler 2008). We chose to have students complete their diary tasks individually in class because they were young learners who could have doubts clarified by the teacher as they emerged and we could have a better account of how they reacted to the diary tasks. In addition to being implemented in class and because using diary tasks over a period of time implies motivated students (Schmitz and Perels 2011), we asked students and teachers in individual interviews (among

other aspects mentioned in the following section) how students had reacted to and had felt about the diary task after the intervention (illustrative student responses in Appendix A). We only used the information from these interviews to give us an idea of how students reacted to the diary task. The two initial applications of the diary task (the first week) were done so that the students could be familiarized with the instrument and the training in how to regulate one's learning. Hence, of the 12 diary tasks each student filled in, only 10 diary task entries per student were considered for analysis in this study. The students ($n=100$) were divided into an experimental group ($n=40$), which was composed of three groups of students and their respective teachers (2) and a control group ($n=60$) which was composed of four groups of students and their respective teachers (4). The assignment of classes to the experimental and control groups was based on whether the teachers were willing to participate in the training sessions. Students in the experimental group were given training in how to regulate one's learning twice a week for a period of 6 weeks in 45 minute sessions by their teacher. All students considered all of the questions in the diary task in each session and filled in the diary task autonomously according to their perceptions of what they had done in each session.

The responses to the open-ended questions of the diary task were coded by two different raters across the 100 participants. We found no example of irrelevant comments (coded as 1) in the students' diary tasks for the first open-ended question (intentions to learn). Examples of students' responses) to the first open-ended question include "today in my class I want to learn to correct my mistakes by reading the story many times" (coded as 9: specific comment related to autonomous goal about strategies to adopt to learn content). For the second open-ended question (anticipations of learning outcomes/performance), examples of students' responses include "Today in my class I will be able to do nothing" (coded as 1=irrelevant comment) and "Today in my class I will be able to evaluate my work by reading my story many times" (coded as 7=specific comment about anticipation related to strategies to adopt to learn content). As for the third open-ended question (self-examination), examples of students' responses include "Today in my class I learned more of nothing" (coded as 1=irrelevant comment) and "Today in my class I learned to evaluate my colleague by listening to his narration a few times" (coded as 7=specific comment relating to cognitive and or affective reflections about strategies used to learn content). For the complete rating scale of the open-ended questions and corresponding examples see Appendices B and C.

In order to verify the inter-rater reliability, we computed an intraclass correlation, which gave us good values for the ICC $(2, 2)=.99$, according to the literature (McGraw and Wong 1996). Essentially, 99 % of the variance in the mean of both raters was true score variance. Furthermore, as a way of measuring academic performance in EFL, students also executed an oral task and an online vocabulary task before and after the intervention.

All teachers were informed about how the diary task should be filled in, as well as how this tool may bring advantages (i.e., awareness of strategic action) and disadvantages (i.e., students getting tired of filling in the diaries), as recommended in the literature (Glogger et al. 2012). The teachers of the experimental group had prior preparation regarding the regulation of learning. They were introduced to Zimmerman's (2000) model of self-regulated learning and to the strategies presented in Zimmerman and Martinez-Pons (1986) study. The researchers discussed these studies with the teachers and answered any doubts that arose. The teachers then had a workshop where they shared how they regulated their own learning with the researchers and whether and how they usually taught learning strategies to their students.

Presenting the training The training in how to regulate one's learning was designed to develop awareness in using strategies to regulate learning during regular EFL classes according to Zimmerman's model of self-regulated learning (2008). Hence, a two-class introduction to the

regulation of learning and to the diary tasks was done and a unit of ten lessons about sports was delivered to the students by their corresponding teacher. Teachers had a daily meeting with the researchers to discuss and prepare classes and to provide feedback regarding the students' behavior during the lessons and the completion of the diary task. The training sessions were observed by a researcher in order to understand whether the training was being implemented properly. Also, in the interviews with students and teachers at the end of the training sessions, we tried to understand their perspective of the occurrences in class. Data from the daily meetings with the teachers, observation notes and interviews were rated by 3 independent raters in order to establish the fidelity of the implementation. This fidelity measure consisted of five questions concerning the objectives of the training regarding the cyclical phases and strategies of self-regulated learning proposed by Zimmerman (2000). Each question was rated on a scale of 0 (was not implemented according to the objectives) to 10 (was fully implemented according to the objectives). Then, in order to verify the inter-rater reliability, we computed an intraclass correlation, which yielded good values for the ICC (2, 2)=.91, according to the literature (McGraw and Wong 1996). Specifically, 91 % of the variance in the mean of both raters was true score variance.

Because of space limitations, we briefly present a synopsis of the training sessions of the experimental group. Note that the main topic of the sessions was sports. In Portugal there are few differences as to the knowledge both men and women have about different sports and inclusively in later years, a substantial amount of effort has been made to provide equal opportunities for both men and women, although the latter still strive to have the same working conditions (Jaeger et al. 2010). Hence, we did not feel that boys and girls would react differently to the topic. The sports were presented by themes (i.e., water sports in lessons three and four, indoor sports in lessons five and six, velocity sports in lessons seven and eight, bravery sports in lessons nine and ten and hit and kick sports in lessons 11 and 12). In each session students had the freedom to choose from two different sports to learn fluency and develop strategies to regulate their learning. Furthermore, the characters in the exercises prompted students to check whether they were proceeding correctly and whether they had to correct their work throughout the lessons. Other prompts, such as attention notes were also included. This would help students be aware of strategies they could use or reconsider when performing tasks. All of the prompts and instructions were provided to students in both English and Portuguese, as recommended by the national primary education curriculum of EFL in Portugal. All content-related texts were presented in English only. Also, the students filled in the first part of the diary task in the first 5 minutes of each session and the second part of the diary task immediately after each session.

Sessions 1 and 2: Introduction to self-regulated learning and diary tasks In the first two sessions students were introduced to the regulation of learning as an autonomous, but guided way of using strategies to learn. They were also told that they could work collaboratively if they wanted to and that they would get prompts that would also guide them in using strategies and in working either individually and in pairs/groups. Explanation of concepts and examples given were simplified considering the population encompassed fourth-grade students. Also, an introduction to how the diary task could be filled in was given. Students were informed that they were going to fill in this diary task at the beginning and end of each session. In these first two sessions students practiced using the diary task so they could understand what they needed to do and the options they had within the diary task. This would allow them to clarify any initial doubts.

Sessions 3 and 4: planning phase: task analysis - strategic planning and setting objectives to practice water sports In session 3 three different images were presented to the students, each

one depicting a different concept (objectives, following instructions and preparing work - planning). Each concept was explained to the student and doubts were clarified. These images were then hung on the wall and remained there for the following 9 lessons. Then, two short stories were introduced to the students about a character practicing water sports so that they could read aloud (narrate). After students read the story aloud, they did a post-reading activity that focused on identifying the character's objectives and the decisions he made in order to prepare for his sport. They also made a plan to help this character achieve his objectives. This exercise also focused on reading practice.

In session 4 the three different images depicting the different concepts (objectives, following instructions and preparing work - planning) were reviewed. Students responded to concept questions in order to remember what they had learned in the previous session. The teacher then asked students to think about their plan from their previous lesson in order to help the story's character perform better in that particular sport. Then, the students had to imagine they were going to practice that sport and write about how they would prepare for it. Lastly, an open class discussion was held so students could reflect as a group about their plans.

Sessions 5 and 6: planning phase: self-motivation beliefs, self-efficacy and task interest in indoor sports In session 5 the concepts from the previous two sessions were reviewed and three different images were presented, each one depicting a different concept (likes to…, can… and responsibility). They presented each concept to students and answered questions. These images were hung on the wall next to the other concept images. Two stories relating to the concepts that were discussed and to indoor sports (i.e. basketball and ping pong) were presented to the students. Then, students worked on an activity where they identified the character's objectives and decisions in order to prepare for the game. They also focused on what the character thought he could do and if he liked to play that particular sport. Students practiced listening, reading and speaking skills.

In session 6 the concepts from the previous two sessions were reviewed and students' doubts were clarified. Then, the students were introduced to making a critique. From here, students proceeded to plan and elaborate a critique of the story (collaboratively or individually). Essentially, they had to summarize the story and express what they liked and did not like about it. In the end, students had to write down whether they had enjoyed developing a critique and why.

Sessions 7 and 8: monitoring phase: self-observation/control and attention focusing in velocity sports In session 7 past concept images were reviewed and three new images were presented, each one depicting a different concept (attention/concentration; revising one's own work; knowing new ways of learning). Each concept was explained to the students and any questions they had were clarified. These images were then hung on the wall next to the other concept images. Two stories relating to the concepts that were discussed, as well as the topic on velocity sports, namely, car racing and cycling, were introduced to students. The activity following the reading aloud exercise focused on identifying the character's objectives and the decisions he made in order to prepare for his sport. It also focused on what that character thought he could do, as well as what he liked to do. Lastly, this activity highlighted aspects dealing with how the character concentrated and corrected his mistakes. After this activity, there was a class discussion about how the character could have done things differently and how the students could find alternative strategies to win the game. The discussion also focused on how students could find alternative strategies when they were experiencing difficulties.

In session 8 the concepts from the previous two sessions were reviewed and any questions students had were clarified. Two concepts, namely "obstacles" and "strategies" were explicitly

focused on. The teachers answered students' questions and introduced a true and false activity on velocity sports, namely, car racing and cycling. This activity focused on obstacles, and on using strategies to overcome them. This exercises also focused on remembering strategies. Then, students interviewed a colleague about their interest in these sports, whether they got distracted when they practiced them and what they did in order to concentrate. Students had to evaluate their interview together (i.e., whether it had been pronounced properly or not). Lastly, these questions were asked about their English class.

Sessions 9 and 10: self-evaluation: self-judgment and self-evaluation in bravery sports In session 9 the concepts from the previous two sessions were reviewed and any questions students had were clarified. The topic on self-evaluation was introduced explicitly. Students' questions were answered. Then, two short stories on bravery sports were presented, namely, fencing and judo, which the students could choose from. The activity following the reading aloud was a multiple choice exercise focusing on identifying the character's objectives and the decisions he made in order to prepare for his sport. It also focused on whether the character thought he could do and liked to do. This activity also highlighted aspects dealing with how the character kept focused, and how he told himself to keep going. Lastly, the exercise dealt with how the character evaluated the outcomes of his combat.

In session 10 the concepts from the previous two sessions were reviewed and any questions students had were clarified. In this lesson, students were instructed to create a story, which would allow them to focus on how to approach an activity, self-motivation beliefs, attention focusing and self-instruction, and lastly, self-evaluation and causal attribution. The story had to be about two children practicing bravery sports and the steps they had to take in order to overcome difficulties and win the match, as well as how they would celebrate if they won. Lastly, students had to reflect about what they liked/did not like about their story and their colleague's story (as they had read to each other) or, if they chose to work collaboratively, they had to proceed in the same manner, but with their partner.

Sessions 11 and 12: self-evaluation: self-reaction - defense/adaptive mechanisms and self-consequence in hit and kick sports In session 11, the concepts from the previous two sessions were reviewed and students' doubts were clarified. The teachers then introduced two stories about hit and kick sports, namely, football and tennis, which students chose to read aloud (narrate) individually or collaboratively. The post-reading activity was a multiple choice activity which reviewed all of the strategy concepts that were reviewed about how to regulate one's learning, including responsibility over outcomes, as well as the vocabulary on these sports. At the end, students spoke in an open-class discussion about how they did in the activity.

In session 12, the concepts from the previous two sessions were reviewed and students' doubts were clarified. The students then had to review the stories from the previous class and had to narrate it together with a colleague. However, before reading it again, they had to recall what had happened in the story. Then, they had to practice reading the story alone until they thought they read it fluently. Hence, they had to write down a strategy plan of what they were going to do in order to get a perfect result. After reading it, they had to evaluate whether they were ready to share with a colleague by answering questions such as: "Do I have to concentrate more?"; "Am I following my plan?"; "Do I have to ask for help?"; "Do I have to read louder?"; "Did I speak well about the sport?"; "Am I happy with my narration?"; "Do I want to read it again before showing my colleague?" The next step was to provide feedback on the narration with other colleagues. The question "What could your colleague have done differently?" guided the conversation between students.

Data analysis of the main study

Process data We used responses from the three dimensions of questions (forethought, performance and self-reflection) from the diary task as our process data. Forethought, Performance and Self-reflection were measured from 1 to 5 on all items in the database. For example, 1="I did not plan my work in class today."; 2="I planned in my class today but did not like to plan it at all"; 3="I liked to plan my work a little"; 4="I liked to plan my work enough"; and 5="I liked to plan my work a lot". The same process was adopted for all variables ("I found it difficult to…", "I made an effort to…" and "I was able to…") in all three blocks (forethought, performance and self-reflection). Then, the responses to the item "I found it difficult to…" were reverse scored and all of the item responses were aggregated by dimensions of forethought, performance and self-reflection, as indicated in the results of the exploratory factor analysis of our pilot study 1 and the confirmatory factor analysis of our pilot study 2. The aggregation was done by day so that we had a mean score of each group for each phase of the self-regulated learning cycle (each dimension) for each day of the training. We used Multilevel Linear Modeling (IBM, SPSS, 22.0) for repeated measures designs in order to measure the difference between the experimental group and the control group concerning perceptions of how they planned, monitored and evaluated their own work throughout the ten sessions. For the analysis, a sample size of 1,000 diary task entries (10 diary task entries per student) was used for each dimension (forethought, performance and self-reflection) at level 1 and of 100 students at level 2. Students were measured on ten occasions.

Each dimension (forethought, performance and self-reflection) constituted a dependent variable, whereas time and the training were considered the covariates. Data was structured at the within-person in time level (level 1) and the between person level (level 2). In terms of our hypothesis formulation, we considered, as recommended in the literature (Cleary 2011; Zimmerman 2000) all three dimensions (phases) as self-regulated learning activity. We used Maximum Likelihood estimation for all analyses, since it is a technique commonly used for large scale samples which provides asymptotically unbiased estimates (McCoach 2010). Variables were introduced in SPSS in three steps so as to test the interaction effects.

Firstly, we examined the intercept-only model to determine how much variability was present in reported self-regulated learning activity at each level. This model may be represented as:

$$\Upsilon_{ti} = \pi_{oi} + \varepsilon_{ti} + u_{oi}$$

While the Υ_{ti} is the observed condition at time t for individual i, the π_{0i} is the intercept (average/grand-mean intercept for reported self-regulated learning activity across students). Lastly, the ε_{ti} corresponds to the variation (estimated errors) in estimating growth within individuals, while u_{0i} is the variation in estimating growth between individual. In this intercept-only model, we used a scaled identity covariance structure for the repeated measures diary task effect and a variance components covariance structure for the intercept random effect because we wanted to examine the amount of variance in the outcome within and between individuals. The scaled identity covariance structure has one estimated parameter and assumes that there is a constant variance across occasions with no correlation between components (Heck et al. 2013).

Secondly, we focused on defining the shape of the growth trajectory. We tested a model including a quadratic time variable and another with orthogonal polynomials, which did not yield any significant results in explaining student growth in reported self-regulated learning activity. The model with linear time did, in fact, yield significant results. Hence, we opted for the linear trend. Thus, the tested model we present next includes four

parameters, namely the intercept and the linear time-related variable as a fixed effect, one random intercept and one residual.

$$\Upsilon_{ti} = \beta_{00} + \beta_{10}a_{ti} + u_{0i} + \varepsilon_{ti}$$

In this model Υ_{ti} corresponds to the observed condition at time t for individual i, β_{00} is the intercept depicting the average initial status mean between individuals, $\beta_{10}a_{ti}$ represents the linear time-related component, u_{0i} is the level 2 random component related to describing any differences in average reported self-regulated learning activity between students, and ε_{ti} corresponds to the errors in predicting the average reported self-regulated learning activity for students. For reported forethought, performance and self-reflection activities, we used a scaled identity covariance structure for the repeated measures diary task effect, which is a simplified covariance structure with only one estimated parameter, and a variance components covariance structure for the intercept random effect (Heck et al. 2013).

Thirdly, because we wanted to understand whether the treatment (training) was related to different growth patterns, we studied any differences in development between the two groups of students. Specifically, we wanted to understand if the change was the same or different for the experimental and control in their reported self-regulated learning activity over time. In order to understand if students in the experimental group reported self-regulated learning activity over time differently from students in the control group, we combined the level 1 model with time specified as linear only to describe students' growth over time, assuming that the intercept varies between subjects and that the time slope is randomly varying. This combined model may be represented as:

$$\Upsilon_{ti} = \beta_{00} + \beta_{01}training_i + \beta_{10}time_{ti} + \beta_{11}time_{ti} * training_i + u_{1i}time_{ti} + u_{0i} + \varepsilon_{ti}$$

Υ_{ti} is the observed condition at time t for individual i, β_{00} is the intercept showing the average initial status mean between individuals, $\beta_{01}training_i$ represents the training variable (coded 0 for no training and 1 for training), $\beta_{10}time_{ti}$ is the linear time-related component, $\beta_{11}time_{ti}*training_i$ is the interaction parameter included to examine if there are different growth trajectories for the individuals in the two groups. Furthermore, $u_{1i}time_{ti}$ and u_{0i} are the level 2 random components related to describing any differences in average reported self-regulated learning activity between students. Lastly, ε_{ti} corresponds to the errors in predicting the average reported self-regulated learning activity for students. We used a scaled identity covariance structure for level 1 and a variance component covariance structure for level 2 of the reported forethought, performance and self-reflection activities, which are simplified covariance structures with only one estimated parameter each (Heck et al. 2013). We wanted our model to be as parsimonious as possible.

The improvement of each model over the previous one was assessed with the corresponding likelihood ratios. This difference in likelihood approximates is in accordance with the chi-square distribution (change in degrees of freedom between models: subtraction of the number of new parameters added to the model from the parameters of the previous model). Thus, we report the differences in the deviances (by subtraction) as evidence that the model with the covariates fits the data better than the model with the intercept and time, and that this latter model fits the data better than the intercept only model.

Pre-post group comparisons We applied group pre-post comparisons in order to test the effects of the training of with respect to the variables: students' intentions to learn, anticipations of learning outcomes and self-examination. We used students' responses from the first learning session (not the introduction sessions) to measure the pre-test and students' responses

from the last learning session to measure the posttest. As mentioned in the description of the instrument, students' answers were coded in order to proceed with a quantitative analysis. Then, we tested the differences between the two groups with analysis of variance (ANOVA) with time as a repeated measurement factor. Because of pretest differences for the variable self-examination between the two groups, we also computed analyses of covariance (ANCOVA) with the pretest value as covariate. Moreover, in order to provide a more detailed analysis as a warrant of this qualitative data (students' responses), we calculated the frequencies of the types of responses students wrote and present examples that may be seen in Appendices B and C. In order to test performance differences between the experimental and control group, we analyzed the mean differences from the oral task and the vocabulary task that were administered before and after the training. This was also calculated with an ANOVA using time as a repeated measurement factor.

Results

Pilot study 1

In the exploratory factor analysis of our first pilot study, we used IBM SPSS 22.0 and FACTOR 9.2 to interpret the internal structure of the diary task with the data gathered from the 78 students. Table 1 shows the correlations among all variables and the descriptive statistics. All of the variables showed skewness values less than 2 and kurtosis values less than 5 as recommended by the literature (Bollen and Long 1993). The data was tested with the Kaiser-Meyer-Olkin (KMO) and the Bartlett's Test of Sphericity for its underlying structure. The KMO measure of sampling adequacy was a reasonable .70, whereas the Bartlett Sphericity was $\chi^2_{66}=$ 1003.84 ($p<.001$), demonstrating that the variables were suitable for factor analyses. According to Bollen and Long (1993), there is multivariate normality if Mardia's coefficient is lower than P (P+2), where P is the number of observed variables. In this study, 12 observed variables were used with a Mardia's coefficient for skewness of $109<12(12+2)=168$ and for kurtosis of $269>$ $12(12+2)=168$. Thus, considering our kurtosis values, we used Unweighted Least Squares (ULS) as the method for factor extraction, an estimation method that does not depend on distributional assumptions (Joreskog 1977). In order to ascertain the appropriate number of factors to retain, various factor retention criteria were applied, specifically, Velicer's MAP test and Horn Parallel analyses. We applied Velicer's MAP test and Horn Parallel analyses because they perform optimally in determining the number of factors to extract (Bandalos and Finney 2010). By using alternative methods of extraction, we intended to propose an approximation to a simple interpretable structure (see Table 1). There were no items with loadings greater than .40 on two or more components, hence, we considered all items with structure coefficients values above .30 (Ford et al. 1986; Bandalos and Finney 2010). Although we tested both oblique (direct oblimin) and orthogonal rotations, we opted for an orthogonal rotation, such as Varimax because we expected the factors to be independent. This preliminary study of the diary task's structure and application suggested a three factor model of forethought ($\alpha=.90$), performance ($\alpha=.88$), and self-reflection ($\alpha=.89$), with good reliability scores according to the psychometric literature (Nunnally 1978). What's more, the values of goodness-of-fit (*GFI*=.96), residuals statistics (*RMSR*=.09) and the Guttman-Cronbach's alpha coefficient ($\alpha=.87$) were good in accordance with the literature (McDonald 1999; Nunnally 1978; Velicer 1976). Hence, the items were grouped into three categories in accordance with Zimmerman's phases of self-regulated learning (2000), namely, forethought, performance and self-reflection actions that had been adopted in class by students.

Table 1 Pilot study 1 item descriptive statistics, exploratory factor analysis parameters, reliability and correlations

Variable	Structure Coefficients			Mean (SD)	Correlations										
	Forethought	Performance	Self-reflection		1	2	3	4	5	6	7	8	9	10	11
I liked to prepare my work.	0.99			4.36 (1.27)											
I found it difficult to prepare my work.	0.45			2.18 (0.73)	.498**										
I made an effort to prepare my work.	0.85			4.12 (1.28)	.880**	.362**									
I was able to prepare my work	0.94			4.21 (1.25)	.935**	.468**	.862**								
I liked to correct my work as I did it.		0.91		4.17 (1.37)	.313**	.073	.371**	.373**							
I found it difficult to correct my work as I did it.		0.36		2.36 (0.96)	.084	.402**	-.002	.078	.385**						
I made an effort to correct my work as I did it.		0.89		4.05 (1.34)	.208	-.049	.355**	.286*	.825**	.335**					
I was able to correct my work as I did it		0.93		4.17 (1.31)	.350**	.103	.372**	.461**	.886**	.413**	.846**				
I liked to evaluate my work			0.89	4.18 (1.40)	.318**	.132	.383**	.356**	.362**	.115	.352**	.336**			
I found it difficult to evaluate my work.			0.49	2.26 (0.90)	-.036	.498**	-.015	-.001	.028	.503**	.053	.040	.424**		
I made an effort to evaluate my work.			0.85	4.03 (1.35)	.279*	.126	.429**	.318*	.367**	.161	.510**	.398**	.870**	.376**	
I was able to evaluate my work.			0.90	4.09 (1.39)	.251*	.111	.304**	.308*	.270*	.024	.322**	.324**	.899**	.403**	.875**
Eigenvalues	4.92	1.93	1.70												
% of Explained Variance	41 %	16 %	14 %												
Cronbach's Alpha	.90	.88	.89												

$^{\dagger} p<0.10$; $^{*} p<0.05$; $^{**} p<0.01$

Pilot study 2

For the confirmatory factor analysis of our second pilot study, we used (IBM, SPSS AMOS 19.0) estimation procedures of unweighted least squares, namely, fit indices such as chi-square, Root Mean Square Error of Approximation (RMSEA), Standardized Root Mean Square Residual (SRMR), Comparative Fit Indices (CFI), Incremental Fit Index (IFI) and Akaike Information Criterion (AIC). The CFI and IFI values close to one indicate a good statistical fit (Bentler 1990), while RMSEA indicates a good fit if equal or less than .08 (Browne and Cudeck 1993). As for the AIC, the lower the value, the better the fit. Lastly, the SRMR should be close to zero for a good fit. We tested various possible models so as to confirm the initial structure of the diary task suggested by the EFA (see Table 2). We attempted to test a model with three factors and with no covariances between the items' error terms (model one); a model with four factors with no covariances (model two); a model with three factors and with covariances (model three) and a model with four factors and also with covariances (model four). For models 3 and 4, since we had identified from the students' responses during the initial development of the diary task that there were items that tapped on the same common key-processes (i.e. liking to do something, finding something difficult, making an effort and being able to accomplish something), we established covariances between the error terms of the items with the same common key-processes. Models one, two and three converged, but model four did not. From the results presented, we chose model three which presented better fit indices. According to the literature (Hooper et al. 2008), the three factor model we opted for presented good reference values [χ^2 (39)=30.93, p=.818, χ^2/df=.793, CFI=1.00, IFI=1.04, RMSEA=.000, LO=.000, HI=.049, SRMR=.054, AIC= 108.930] in comparison with the competing models. Most of the relationships between each factor and corresponding items were higher than .5, as suggested in the literature as a good cut-off (e.g., Bandalos and Finney 2010). All unstandardized path coefficients were significant at p<.05. What's more, the construct reliability scores were higher than .80 (Hair et al. 2010), while the Average Variance Extracted (AVE) scores were higher than .50, therefore supporting convergent validity (Henseler et al. 2009). The variables' discriminant validity was also confirmed with all of the Average Shared Variance scores below the AVE scores (Hair et al. 2010). From these results, we maintained the structure of the diary task in accordance with Zimmerman's phases of self-regulated learning (2000) and as the EFA had initially proposed.

Main study

Prior to testing our hypotheses, we computed the means, correlations and reliability coefficients of each variable (Table 3).

Test of hypotheses For hypothesis 1 we had expected that there would be different growth rates of reported self-regulated learning activity over time between the students who had experienced the training and the students who had not. Again, by self-regulated learning activity we mean forethought, performance and self-reflection activities (the three phases of Zimmerman's self-regulated learning model). Table 4 presents the model fit information (likelihood ratios) and estimates for the fixed and random effects of all three models.

At level 1, the variance corresponds to the variability in the average students' reported self-regulated learning activity estimates around their own growth trajectory (Singer & Willet, 2003). The estimates of variance for levels 1 and 2 of reported forethought activity (Z_w=6.072, p<.001), reported performance activity (Z_w=5.796, p<.001) and reported self-reflection activity (Z_w=4.988, p<.001) suggest that there was sufficient variation in intercepts across

Table 2 Fit indices of tested models and descriptive statistics, path coefficients, Construct Reliability, Average Variance Extracted, and Average Shared Variance of Model 3

CFA Models	χ2	df	Sig.	χ2/df	CFI	IFI	RMSEA	AIC	SRMR
Model 1	450.0	51	.000	8.825	.616	.626	.307	528.07	.117
Model 2	388.2	48	.000	8.089	.673	.682	.292	472.27	.090
Model 3	30.93	39	.818	0.793	1.00	1.04	.000	108.93	.054
Factors		Mean (SD)	1	2	3				
1. Forethought		3.21 (1.07)							
2. Performance		3.15 (1.07)	.74**						
3. Self-reflection		2.99 (1.14)	.52**	.37**					
Path Coefficients [Min; Max]			[.50;.91]	[.54; .94]	[.59; .91]				
CR			.88	.90	.88				
AVE			.66	.69	.66				
MSV			.58	.58	.25				
ASV			.41	.35	.18				
Cronbach's Alpha			.86	.88	.86				

For the covariances between errors we followed the AMOS posttest analyses and co-varied items sharing the same content. *CR*, construct reliability; *AVE*, average variance extracted; and *ASV*, average shared variance

Table 3 Descriptive statistics. Reliability coefficients. and correlations of variables for multilevel analysis

Variables	Correlations				Group	Level 1 (N=1,000)										Level 2 (N=100)
	1	2	3	4		0	1	2	3	4	5	6	7	8	9	
						M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)	M (SD)
1. SRL Training		.51**	.54**	.31**												
2. Forethought	.34**	(.86)	.89**	.52**	EG	3.90 (.71)	4.10 (.65)	4.11 (.62)	4.20 (.41)	3.98 (.77)	4.03 (.76)	4.10 (.68)	4.16 (.45)	4.08 (.67)	4.00 (.83)	4.07 (.34)
					CG	3.98 (.66)	3.90 (.64)	3.05 (.84)	2.89 (.95)	3.39 (.75)	3.92 (.95)	3.97 (.58)	3.02 (.91)	2.97 (.95)	3.44 (.76)	3.45 (.58)
3. Performance	.33**	.63**	(.81)	.43**	EG	3.40 (1.0)	3.85 (.85)	4.13 (.52)	4.16 (.45)	3.91 (.88)	4.01 (.77)	4.17 (.55)	4.23 (.42)	4.12 (.63)	4.20 (.65)	4.01 (.41)
					CG	3.97 (.75)	3.37 (.90)	2.95 (.85)	2.98 (.96)	3.33 (.81)	3.88 (1.0)	3.92 (.61)	3.06 (.88)	2.95 (.94)	3.47 (.69)	3.39 (.25)
4. Self-reflection	.16**	.46**	.40**	(.92)	EG	3.54 (1.0)	3.71 (.98)	3.59 (1.0)	3.67 (.99)	3.34 (1.0)	3.41 (1.0)	3.85 (.83)	3.11 (1.0)	3.66 (1.0)	3.80 (.97)	3.57 (.62)
					CG	3.80 (.99)	3.63 (.82)	2.90 (.77)	2.79 (.88)	3.07 (.80)	3.63 (1.1)	3.87 (.73)	2.96 (.88)	2.45 (.56)	3.25 (.81)	3.24 (.40)

[†] $p<0.10$; [*] $p<0.05$; [**] $p<0.01$, *EG*, experimental group; *CG*, control group. Correlations below the diagonal are day level correlations ($N=1,000$). Correlations above the diagonal are person-level correlations ($N=100$). Cronbach's alpha coefficients are reported in brackets on the diagonal. The Level 1 means and standard deviations are reported according to the time variable (from 0 to 9)

Table 4 Fixed and random effects parameter estimates for models predicting reported self-regulated learning activity

Parameter	Reported Self-Regulated Learning Activity								
	Forethought			Performance			Self-reflection		
	Intercept-only	Intercept+Time	With Predictors	Intercept-only	Intercept+Time	With Predictors	Intercept-only	Intercept+Time	With Predictors
Fixed Effects									
Intercept	3.70**(.05)	3.82**(.07)	3.28**(.16)	3.64**(.05)	3.54**(.06)	3.22**(.16)	3.37**(.05)	3.52**(.07)	3.41**(.19)
Time		−0.02**(.00)	−0.09**(.02)		0.00**(.00)	−0.09**(.03)		−0.03**(.01)	−0.11**(.02)
Training			0.38***(.11)			0.26**(.11)			0.08 (.13)
Time*Training			0.05**(.02)			0.07**(.02)			0.06**(.02)
Random Effects									
Repeated measures	0.47**(.02)	0.47***(.02)	0.44***(.02)	0.57**(.03)	0.57**(.03)	0.53**(.02)	0.80**(.03)	0.79**(.04)	0.78**(.04)
Intercept	0.29**(.04)	0.29**(.04)	0.13**(.03)	0.26**(.04)	0.26**(.04)	0.11**(.03)	0.19**(.03)	0.19**(.04)	0.16**(.03)
Time			0.00**(00)			0.00 (.00)			0.00 (.00)
Deviance	2299.17	2286.92	2232.76	2457.97	2449.64	2383.55	2741.46	2730.81	2712.41
AIC	2305.17	2294.92	2246.76	2463.97	2457.64	2397.55	2747.46	2738.81	2726.41
BIC	2319.89	2314.55	2249.68	2478.69	2477.27	2431.90	2762.18	2758.44	2760.76

Standard errors are in brackets. † $p<0.10$; * $p<0.05$; ** $p<0.01$. The extended values of the random effects of the Time variable in the models with predictors are .002** for forethought, .003** for performance and .0004 for self-reflection. The extended value of the fixed effect of the Time variable in the Intercept+Time model for Performance is 0.002**

students. We estimated the proportion of variance (ICC) using a one-tailed test for variances, giving us 38 % of variance between individuals and 62 % of variance within individuals for forethought. As for performance, we estimated 31 % between individuals and 69 % within individuals, while 19 % between individuals and 81 % within individuals for self-reflection. Thus, we concluded that there was variance within and between students' reported self-regulated learning activity over time. The intercept-only model, which included only the intercept, was compared to intercept+time model. The intercept+time model displayed a significant improvement over the intercept-only model (forethought: *deviance*=12.25, *df*=1, *p*<.01; performance: *deviance*=8.33, *df*=1, *p*<.01; self-reflection: *deviance*=10.65, *df*=1, *p*<.01). In this second model, the intercept corresponds to the students' reported self-regulated learning activity at the beginning of the study. The linear time variable was significant in explaining student growth in reported self-regulated learning activity (forethought, performance and self-reflection).

The model containing the predictor variables and the interaction between them presented a significant improvement over the intercept+time model (forethought: *deviance*=54.16, *df*=3, *p*<.01; performance: *deviance*=66.09, *df*=3, *p*<.05; self-reflection: *deviance*=18.40, *df*=3, *p*<.01). The results from this third model indicate that the students in the control group began with a mean of 3.28 for reported forethought activity, 3.22 for reported performance activity, and 3.41 for reported self-reflection activity. Results also show that there was an initial significant difference between groups for reported forethought activity and reported performance activity, as the experimental group started off with 3.66 for the first and with 3.48 for the latter. However, there was no initial significant difference between groups for reported self-reflection activity. Over each interval the scores of the control group decreased on average by -.09 points for reported forethought activity, by -.09 points for reported performance activity, and by -.11 points for reported self-reflection. Moreover, our results reveal that the experimental group decreased less than the control group. On average, the experimental group decreased by .04 for reported forethought activity (i.e. Fig. 3), by .02 for reported performance activity (i.e. Fig. 4) and by .05 for reported self-reflection activity (i.e. Fig. 5). As we can see in Figs. 3, 4 and 5, the effect of the training (experimental group) was in reducing the negative change over time seen in the control group.
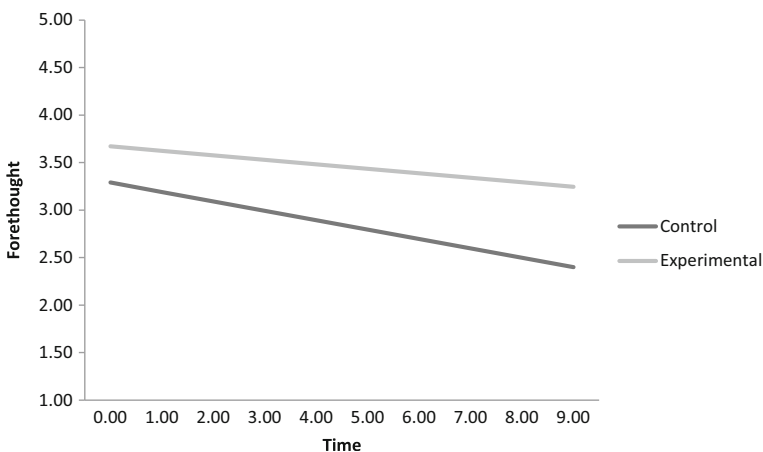

Fig. 3 Fitted trajectories of the experimental and control group for reported forethought activities
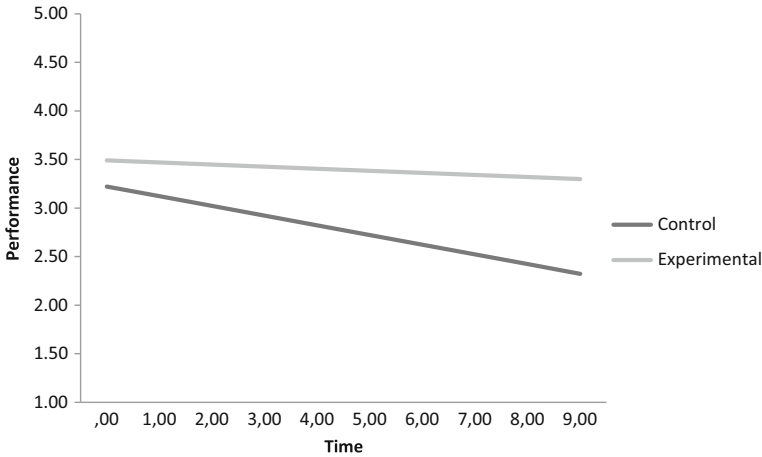
Fig. 4 Fitted trajectories of the experimental and control group for reported performance activities

Furthermore, our results suggest that there were significant differences in the growth rates of reported forethought, performance and self-reflection activities over time between the two groups. That is, over each interval, there was a difference of .05 points for reported forethought, of .07 for reported performance and of .06 for reported self-reflection. In sum, while the control group significantly decreased in reported self-regulated learning activity (forethought, performance and self-reflection) the experimental group managed to decline less. Thus, we can conclude that there were different growth rates of reported self-regulated learning activity over time between students who experienced training and students who did not. (H1).

Hypothesis 2 stated that the students who experienced training in how to regulate one's learning would report their reflections more autonomously and specifically in their diary task than students who did not. There were no significant differences between the two groups in the pretests for intentions to learn [t (98)=−1.68, p>.05], and for anticipations of learning outcomes [t (98)=−0.19, p>.05]. There were significant initial differences between the groups in the pretests for self-examination [t (98)=−2.27, p<.05]. We performed an ANCOVA for this
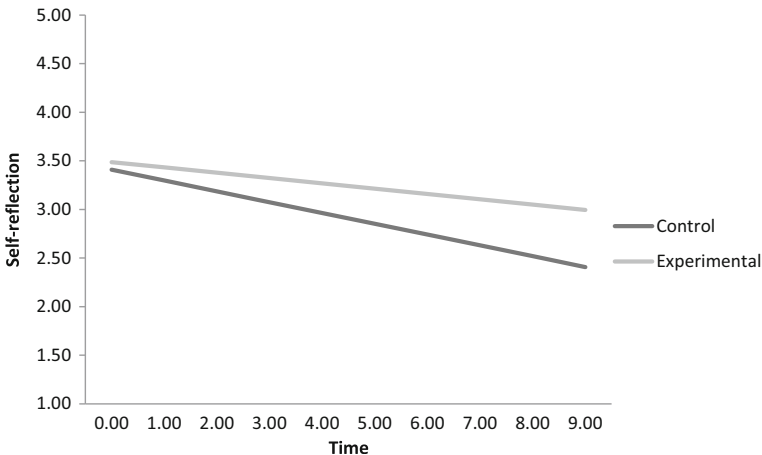

Fig. 5 Fitted trajectories of the experimental and control group for reported self-reflection activities

last variable as mentioned in the method section. Table 5 shows these results, along with the means and standard deviations. In light of Cohen's (1988) distinction between small ($h_p^2 = .01$), medium ($h_p^2 = .06$) and large effects ($h_p^2 = .14$), we found medium to large effects for students' reported reflections (intentions, anticipations and self-examination).

The results revealed significant differences in favor of the experimental group when compared to the control group. Specifically, there was a significant interaction between time and group for students' reflections (i.e. intentions, anticipations and self-examination), having the experimental group scored significantly higher than the control group in the posttest. These results indicate that at the end of the intervention, the students in the experimental group were more autonomous and specific when they wrote about what they intended to do in class (i.e. set goals for themselves). Moreover, these results reveal that the students in the experimental group were more autonomous and specific when they wrote about what they thought they were capable of doing in class. Lastly, these results show that the students in the experimental group were more autonomous and specific when they wrote about the what they had learned, including the appropriateness of their efforts, efficacy, thoughts and actions. Thus, hypothesis 2 was supported, confirming that students who experienced training in how to regulate one's learning reported their reflections more autonomously and specifically in their diary task than students who did not. In order to provide a detailed analysis of students' comments to the open-ended questions, we present in Appendices B and C estimated percentages and examples of the types of comments that students in the experimental group and in the control group wrote in their diary tasks in the pre and posttest. In agreement with the mean differences presented, these percentages showed a shift in the comments of the experimental group towards more autonomous and specific reports of their reflections.

Lastly, we hypothesized (H3) that the students who experienced the training in how to regulate one's learning would have better academic performance in EFL than students who did not. The results from an ANOVA with time as a repeated measurement factor (Table 5) supported this last hypothesis. No initial significant differences were found between groups in the pretest for the oral task [t (98)=−1.79, $p>.05$], and for the vocabulary task [t (98)=−0.28, $p>.05$]. Moreover, there was a significant interaction between time and group for students' performance.

Table 5 Means (standard deviations) and results for the interaction Time x Group regarding intentions to learn, anticipations of learning outcomes and self-examination

DV	Group	Time		df	F	h_p^2
		Pretest M (SD)	Posttest M (SD)			
Intentions	EG	4.20 (1.3)	5.70 (2.0)	1/98	27.75**	.22
	CG	3.66 (1.6)	3.20 (1.1)			
Anticipations	EG	4.47 (.90)	5.83 (.80)	1/98	11.54**	.11
	CG	4.43 (1.1)	4.80 (1.0)			
Self-examination	EG	4.10 (1.6)	6.30 (.60)	1/98	42.60**	.30
	CG	3.36 (1.5)	3.16 (1.3)			
Oral Competence	EG	1.91 (.33)	3.16 (.87)	1/98	65.69**	.40
	CG	1.75 (.47)	1.97 (.35)			
Vocabulary Competence	EG	0.53 (.13)	0.62 (.14)	1/98	15.36**	.14
	CG	0.52 (.10)	0.51 (.15)			

*$p<0.05$. **$p<0.01$; *EG*, experimental group; *CG*, control group. Normality assumptions were tested prior to analysis and fulfilled

The experimental group scored significantly higher in the oral task and in the vocabulary task in the posttest. Hence, the third hypothesis of this study was also confirmed, suggesting that the training in how to regulate one's learning influenced students' academic performance.

Discussion

As we have mentioned in this study, assessing students' perceptions regarding their self-regulated learning experiences in a classroom context is a difficult task which requires measures that are process-oriented (Zimmerman 2008). This study contributes to the existing literature because it presents an instrument that goes beyond typical survey self-report measures to assess students' perceptions of their self-regulated learning processes as they happen (e.g. Azevedo and Cromley 2004; Schmitz 2006; Schmitz et al. 2011). Following a microanalytic methodology, the diary task we used can be a highly effective approach for assessing changes in reported self-regulated learning activity and students' reflective thinking (Cleary 2011). The diary task DOGS-RL can be proximal and specific to any determined academic task, allowing us to track students' reported self-regulated learning activity and reflective thinking throughout a training program in how to regulate one's learning. Furthermore, this study focused on a domain-area (EFL) that is not typically studied in terms of self-regulated learning processes, as recommended in the literature (Greene et al. 2010). This study also gives its contribution to this field in the sense that it used, not only a factor analysis to determine the structure of the DOGS-RL, but also, a process analysis, as suggested in the literature (Cleary 2011; Schmitz et al. 2011), with a multilevel analysis approach for repeated measurements as in other areas that used diary tasks to measure process data (Jett et al. 2010; Rowe et al. 2009; Whitty et al. 2011). Moreover, the quasi-experimental pre-posttest design we used within a classroom context with primary school students and teachers allowed us to test the effects of the training in how to regulate one's learning on academic performance outcomes. What's more, this approach allowed us to validate the diary task DOGS-RL as a monitoring tool for strategic reflection and reporting and for capturing changes in the process of learning.

Our first hypothesis was confirmed with our results from the multilevel analysis of the quantitative items of the diary task. Specifically, we verified that the experimental group's growth rates of reported forethought, performance and self-reflection activities decreased less than those of the control group. Hence, we concluded that there were different growth rates of reported self-regulated learning activity over time between students who experienced training and students who did not, allowing us to answer our first research question affirmatively. The reason why students decreased their reported self-regulated learning activity could be because, and as seen in previous studies, students of this age group may initially overestimate their meta cognitive awareness of how they function in the classroom (Allwood et al. 2008; Lipko-Speed 2013). The fact that the experimental group decreased less could be explained by some of the information the teachers provided us with in the daily meetings. In sum, the teachers believed that the students in the control group needed explicit guidance in how to regulate their learning in order to fill in the diary because they were not used to planning, monitoring and evaluating their work in class on a regular basis, or even at all in an explicit manner. These results are in accordance with previous literature (i.e., Glogger et al. 2012; Perels et al. 2007, Schmitz and Perels 2011) and even meet some of the challenges regarding suggestions for future research regarding the measurement of self-regulated learning, namely with respect to changes in learners' interest, task difficulty and usefulness, and effort over a period of time (Cascallar et al. 2006; Moos and Azevedo 2008). Nonetheless, unlike most studies (i.e., Glogger et al. 2012; Perels et al. 2007, Schmitz and Perels 2011), these results pertain to diary tasks that were

completed in a classroom environment and not sent for homework, which allowed the teachers to help students as doubts emerged and allowed us to better witness how these diary tasks were filled in by the students, as well as how they reacted to it (i.e., doubts). This option to implement the diary task in class had to do mainly with the fact that we worked with primary school students in the fourth grade, unlike most studies that focus on older students. Having students fill in their diary tasks in class allowed us to understand any differences between classroom practices and provided our study with further ecological validity. Considering students were from different classes with a different teacher, the varying classroom practices may have contributed to the fact that the results showed initial significant differences between groups for forethought and performance activities. What's more, in the daily meetings, teachers informed the researchers that while students were required to self-evaluate their work in general at the end of each academic period, they were not used to planning and monitoring their work in an explicit manner. This may have been one of the reasons why there were no initial significant differences between groups for reported self-reflection. Further research into this topic would be needed to fully understand the causes of the initial differences between groups in reported forethought and monitoring activities.

We also hypothesized (H2) that students who experienced training in how to regulate one's learning would report their reflections more autonomously and specifically in their diary task. A comparison between groups with an ANOVA using time as a repeated measurement factor supported this hypothesis, indicating similar findings to what has been investigated in other studies. Specifically, that students who experience training in how to regulate one's learning are autonomous and specific in reporting their reflections rather than descriptive about learning content (Perels et al. 2007; Otto 2007; Schunk and Zimmerman 1998; Stoeger and Ziegler 2008; Stoeger and Ziegler 2011; Wolters et al. 2003; Zimmerman 2000). However, unlike most studies, we provided a detailed analysis of our qualitative data that further supports and warrants our second hypothesis. Specifically, we presented percentages of the types of comments students made, along with examples in order to illustrate exactly what students reported (Appendices B and C). These reports showed that in fact, the perceptions students in the experimental group reported in the open-ended questions shifted more towards strategic actions to learn than the control group in the posttest.

Our last hypothesis stated that students who experienced training in how to regulate one's learning would have better academic performance in EFL. A comparison between groups with an ANOVA using time as a repeated measurement factor confirmed this hypothesis. These results are similar to the results from previous studies, where students who used diary tasks and experienced training in how to regulate one's learning performed better academically (Glogger et al. 2012; Otto 2007; Perels et al. 2009). However, while most of these studies investigated training in how to regulate one's learning and diary tasks in math, we opted to contextualize our study in EFL classes, which we feel is strength of this study because there is less research considering diary tasks and self-regulated learning in this field (Greene et al. 2010).

There could be some limitations to our study concerning internal validity issues typical of designs with repeated measurements, namely a regression threat which encompasses situations where subjects are tested several times and their scores tend to regress towards the mean. Another limitation may be a maturation threat where subjects may change during the course of the experiment. Furthermore, the fact that there were initial significant differences between the groups for reported forethought and performance activities could indicate a pre-existing advantage of the experimental group over the control group in this domain. Moreover, since students worked in different classes with different teachers, these differences could have arisen from the difference in classroom practices, which we did not control in the multilevel analysis. The fact that there was a non-random assignment of classes to treatment, could also constitute

a strong limitation because we were limited to teachers' willingness to participate in the training. What's more, we studied classes from one country only. Hence, it would be interesting for future research to investigate cross-cultural differences in this area. Also, we did not consider any individual cases or individual learning sessions for analysis purposes. We could also consider examining this type of data process with time series analysis. We intend to do so in future studies, where this diary task will be used and other aspects of students' learning process and specific characteristics of this training in how to regulate one's learning could be contemplated with other samples.

Overall, our findings are in accordance with what the literature has suggested about learners being self-regulated in the sense that they are cognitively, metacognitively, motivationally, and behaviorally active participants in their own learning process (Wolters et al. 2003; Zimmerman and Martinez-Pons 1986). Furthermore, if we consider individual reflection regarding one's own actions and thoughts as part of metacognition, this means that metacognitive reasoning is essential for students to engage in planning, monitoring and self-reflection in class. Similar to what has been discussed in previous literature (Cleary 2011), motivation also played a key role in students' perceptions of their planning, monitoring and self-reflection activities because of the interest they revealed throughout the sessions regarding these experiences in their responses.

Furthermore, although we only used (and only present in the Appendices section) the information from students' interviews at the end of the training as a motivational measure to illustrate how students felt about using their diary task, we posit that this step was important and is also a strength of this study. Namely because it allowed us to understand that students felt motivated to do the diary task (see Appendix A). This was an issue that we were concerned about because students had to fill in many diary tasks, which could have at some point, demotivated them, as mentioned in previous studies (Glogger et al. 2012). An interesting direction for future research would also be to understand how students feel about completing diary tasks through applications with different time lengths and whether this has any implications on how students respond to questions in their diary tasks.

We feel that the implications of our results move towards a better understanding of how there can be change in learners' perspectives about how they regulate their learning and how this change can be measured with a diary task. Moreover, this change can also occur in regulation itself, thus, it would be interesting to add to this diary task specific regulation activity that would measure how the learner actually regulates his/her own learning. With this in mind, research and theory should continue to focus on learners' perceptions, but increasingly from a process approach, as we have tried to demonstrate in this study. We believe that this approach is central to learning and to the regulation of learning because, as Bandura stated (2006), individuals' behaviors are influenced by external and internal factors that are in constant change. Hence, it is vital to measure this change and to understand how overt and covert influences affect learners (Zimmerman 2000) by collecting data from the learners' perspective and from the environment. This type of assessment could aid schools and teachers in keeping a track of students' continuous development over time. This itself could have important implications for a personalization of students' evaluations in school (Cleary 2011). What's more, if teachers have information about what their students perceive of the tasks proposed to them and of how they perform and regulate their learning in class, they can mold their pedagogical practices to fit students' needs in terms of cognitive and motivational processes.

Future studies could focus on how teachers could use their pedagogical practice to promote individual reflection about the learning process. They could themselves practice regulating their own learning, co-regulate and share regulation with colleagues in order to better guide

students and teach them learning strategies explicitly. It would also be interesting to measure in future research how training in how to regulate one's learning could be designed with contemporary technological tools to promote and capture self-regulated learning processes, as well as students' perceptions about these same processes (Azevedo and Cromley 2004). More specifically, future research could focus on computational design aspects of training in how to regulate one's learning in various curriculum-based learning environments, as recommended in previous literature (Graesser and McNamara 2010), as well as on an interactive and dialectic online version of the diary task DOGS-RL that could promote these processes and reflections.

Ultimately, the implications of this study are of great importance for researchers and practitioners working with elementary school children. Basically because the diary task presented here can be adopted from early on in first-grade and adapted to the different grades in primary education (i.e., substituting written form with audio form). Considering this diary-task has the potential to promote reflectiveness in students, it would be wise to start early as has been recommended in the literature (Whitebread et al. 2009). It would also be interesting to examine with a longitudinal study how children could develop their reflective thinking throughout the primary school years. Furthermore, given the importance of collaboration in learning (Järvelä and Hadwin 2013), this diary task could be used/adapted and tested as a collaborative diary task as well, where students could practice sharing goals, anticipations and reflections of their collaborative performance and regulation. This could be an opportunity for developing socially-shared regulated learning from early on.

In sum, we argue that this diary task captured change in students' perceptions regarding their experiences of planning, monitoring and evaluating their own work, as other instruments that were tested in the literature (Klug et al. 2011; Schmitz and Wiese 2006). Teachers can use the DOGS-RL in their class with their students in order to understand how they view their self-regulated learning actions in class, which could have important implications, not only for pedagogical classroom practice, but also for how their students reflect about planning, monitoring and self-evaluating. Lastly, using learning diary tasks to capture learners' perceptions of self-regulated learning experiences as they participate in different learning environments also provides educational psychologists with an important tool/method for assessing cognitive and perceptual changes that are difficult to measure with traditional evaluation methods.

References

Allwood, C. M., Innes-Ker, Å., Holmgren, J., & Fredin, G. (2008). Children's and adults' realism in their event-recall confidence in response to free recall and focused questions. *Psychology, Crime & Law, 14*, 529–547.

Alterio, M. (2004). Collaborative journaling as a professional development tool. *Journal of Further and Higher Education, 28*, 321–332.

Azevedo, R., & Cromley, J. (2004). Does training on self-regulated learning facilitate students' learning with hypermedia? *Journal of Educational Psychology, 96*(3), 523–535. doi:10.1037/0022-0663.96.3.523.

Bandalos, D. L., & Finney, S. J. (2010). Exploratory and confirmatory factor analysis. In G. R. Hancock & R. O. Mueller (Eds.), *A reviewer's guide to quantitative methods in the social sciences: Revise, accept, reject* (pp. 93–114). New York: Routledge.

Bandura, A. (2006). Toward a psychology of human agency. *Perspectives on Psychological Science, 1*(2), 164–180. doi:10.1111/j.1745-6916.2006.00011.x.

Bandura, A., & Locke, E. (2003). Negative self-efficacy and goal effects revisited. *Journal of Applied Psychology, 88*(1), 87–99.

Bares, C. B. (2011). Emerging metacognitive processes during childhood: implications for intervention development with children. *Children and Adolescent Social Work, 28*(4), 291–299. doi:10.1007/s10560-011-0233-1.

Bentler, P. M. (1990). Comparative fit indexes in structural models. *Psychological Bulletin, 107*(2), 238–46.

Boekaerts, M. (2002). *Motivation to Learn*. The International Academy of Education and Bureau of Education. Available at http://www.ibe.unesco.org.

Bollen, K. A., & Long, J. S. (1993). *Testing structural equations models*. Newbury Park: Sage.

Boud, D. (2001). Using journal writing to enhance reflective practice. *New Directions for Adult and Continuing Education, 90*, 9–17. doi:10.1002/ace.16.

Browne, M. W., & Cudeck, R. (1993). Alternative ways of assessing model fit. In K. A. Bollen & J. S. Long (Eds.), *Testing structural equation models* (pp. 136–162). Beverly Hills: Sage.

Cascallar, E., Boekaerts, M., & Costigan, T. (2006). Assessment in the evaluation of self-regulation as a process. *Educational Psychology Review, 18*(3), 297–306. doi:10.1007/s10648-006-9023-2.

Cleary, T. (2011). Emergence of self-regulated learning microanalysis. In B. Zimmerman & D. Schunk (Eds.), *Handbook of self-regulation of learning and performance* (pp. 87–101). New York: Routledge.

Council of Europe (2011). *Common European Framework of Reference for Languages: Learning, Teaching, Assessment*. Council of Europe. Retrieved from: http://www.coe.int/t/dg4/linguistic/Cadre1_en.asp

European commission (2013). European Strategic Framework for Education and Training (ET 2020), Language learning at pre-primary school level: making it efficient and sustainable: a policy handbook. Retrieved from http://ec.europa.eu/languages/policy/language-policy/documents/early-language-learning-handbook_en.pdf

Ford, J. K., MacCallum, R. C., & Tait, M. (1986). The application of exploratory factor analysis in applied psychology: critical review and analysis. *Personnel Psychologist, 39*, 292–314.

Gbabremani-Gbajar, S., & Mirbosseini, S. (2005). English class or speaking about everything class? Dialogue journal writing as a critical EEL literacy practice in an Iranian big school. *Language, Culture, and Curriculum, 18*, 286–299.

Glaser, C., & Brunstein, J. C. (2007). Improving fourth-grade students' composition skills: effects of strategy instruction and self-regulation procedures. *Journal of Educational Psychology, 99*(2), 297–310. doi:10.1037/0022-0663.99.2.297.

Glogger, I., Schwonke, R., Holzapfel, L., Nuckles, M., & Renkl, A. (2012). Learning strategies assessed by journal writing: prediction of learning outcomes by quantity, quality, and combinations of learning strategies. *Journal of Educational Psychology, 104*(2), 452–468.

Graesser, A. C., & McNamara, D. S. (2010). Self-regulated learning in learning environments with pedagogical agents that interact in natural language. *Educational Psychologist, 45*, 234–244. doi:10.1080/00461520.2010.515933.

Greene, J. A., Bolick, C. M., & Robertson, J. (2010). Fostering historical knowledge and thinking skills using hypermedia learning environments: the role of self-regulated learning. *Computers & Education, 54*, 230–243. doi:10.1016/j.compedu.2009.08.006.

Gudmundson, E. (2009). Guidelines for translating and adapting psychological instruments. *Nordic Psychology, 61*(2), 29–45. doi:10.1027/1901-2276.61.2.29.

Hair, J., Black, W., Babin, B., & Anderson, R. (2010). *Multivariate data analysis* (7th ed.): Prentice-Hall, Inc. Upper Saddle River, NJ, USA.

Heck, R., Thomas, S., & Tabata, L. (2013). *Multilevel and Longitudinal Modeling with IBM SPSS* (2nd ed.), Routledge.

Henseler, J., Ringle, C. M., & Sinkovics, R. R. (2009). The use of partial least squares path modeling in international marketing. *Advances in International Marketing, 20*, 277–319.

Hiemstra, R. (2001). Uses and benefits of journal writing. In L. M. English & M. A. Gillen (Eds.), *Promoting journal writing in adult education* (pp. 19–26). San Francisco: Jossey-Bass.

Hofer, B. (2005). The legacy and the challenge: Paul Pintrich's contributions to personal epistemology research. *Educational Psychologist, 40*, 95–105.

Hooper, D., Coughlan, J. & Mullen, M. R. (2008). Structural equation modeling: Guidelines for determining model fit. *The Electronic Journal of Business Research Methods, 6* (1), 53–60, available online at www.ejbrm.com.

Jaeger, A., Gomes, P., Silva, P., & Goellner, S. (2010). Trajetórias de mulheres no esporte em Portugal: assimetrias, resistências e possibilidades [Women's trajectories in sports in Portugal: asymmetries, resistance and possibilities]. *Movimento, 16*(1), 245–267.

Järvelä, S., & Hadwin, A. (2013). New frontiers: regulating learning in CSCL. *Educational Psychologist, 48*(1), 25–39. doi:10.1080/00461520.2012.748006.

Jett, S., LaPorte, D., & Wanchisn, J. (2010). Impact of exposure to pro-eating disorder websites on eating behaviour in college women. *European Eating Disorders Review, 18*, 410–416. doi:10.1002/erv.1009.

Joreskog, K. G. (1977). Factor analysis by least-squares and maximum-likelihood methods. In K. Enslein, A. Ralston, & H. S. Wilf (Eds.), *Statistical methods for digital computers* (pp. 125–153). New York: Wiley.

Kaur, S. (2003). Using student journals for evaluating course experience. *UltiBASE Journal, RMIT University, Australia*. Available online at http://ultibase.rmit.edu.au/Articles/nov03/kaur2/htm.

Kember, D. (2004). Interpreting student workload and the factors which shape students' perceptions of their workload. *Studies in Higher Education, 29*(2), 165–184. doi:10.1080/0307507042000190778.

Kember, D., Jones, A., Loke, A., Mckay, J., Sinclair, K., Tse, H., Webb, C., Wong, F., Wong, M., & Yeung, E. (1999). Determining the level of reflective thinking from students' written journals using a coding scheme based on the work of Mezirow. *International Journal of Lifelong Education, 18*(1), 18–30. doi:10.1080/026013799293928.

Kember, D., McKay, J., Sinclair, K., & Wong, F. (2008). A four-category scheme for coding and assessing the level of reflection in written work. *Assessment & Evaluation in Higher Education, 33*(4), 369–379. doi:10.1080/02602930701293355.

Kir, E. (2012). Diary keeping in English lessons. *Journal of New World Sciences Academy, 7*(4), 1082–1094.

Kirsh, C. (2012). Developing children's language learner strategies at primary school. *Education 3–13, 40*(4), 379–399. doi:10.1080/03004279.2012.691372.

Klug, J., Ogrin, S., Keller, S., Ihringer, A., & Schmitz, B. (2011). A plea for self-regulated learning as a process: modelling, measuring and intervening. *Psychological Test and Assessment Modeling, 53*(1), 51–72.

Larkin, S. (2010). *Metacognition in young children*. NY: Routledge.

Lipko-Speed, A. (2013). Can young children be more accurate predictors of their recall performance? *Journal of Experimental Child Psychology, 114*, 357–363. doi:10.1016/j.jecp.2012.

Määttä, E., Järvenoja, H., & Järvelä, S. (2012). Triggers of students' efficacious interaction in collaborative learning situations. *Small Group Research, 43*(4), 497–522. doi:10.1177/1046496412437208.

McCoach, D. B. (2010). Hierarchical linear modeling. In G. R. Hancock & R. O. Mueller (Eds.), *A reviewer's guide to quantitative methods in the social sciences: Revise, accept, reject* (pp. 123–140). Routledge: New York.

McDonald, R. P. (1999). *Test theory: A unified treatment*. Mahwah: LEA.

McGraw, K. O., & Wong, S. P. (1996). Forming inferences about some intraclass correlation coefficients. *Psychological Methods, 1*, 30–46.

McNamara, D. (2011). Measuring deep, reflective comprehension and learning strategies: challenges and successes. *Metacognition and Learning, 6*, 195–203. doi:10.1007/s11409-011-9082-8.

Moos, D. C., & Azevedo, R. (2008). Exploring the fluctuation of motivation and use of self-regulatory processes during learning with hypermedia. *Instructional Science, 36*, 203–231. doi:10.1007/s11251-007-9028-3.

Nunnally, J. C. (1978). *Psychometric theory* (2nd ed.). New York: McGraw-Hill.

Otto, B. (2007). *SELVES – Schüler, Eltern- und Lehrertrainings zur Vermittlung effektiver Selbstregulation [SELVES - student, parent and teacher training for teaching effective self-regulation]*. Berlin: Logos.

Oxford, R. L. (2003). *Language learning styles and strategies: An overview*. Oxford: GALA.

Perels, F., Otto, B., Landmann, M., Hertel, S., & Schmitz, B. (2007). Self-regulation from a process perspective. *Journal of Psychology, 215*(3), 194–204.

Perels, F., Dignath, C., & Schmitz, B. (2009). Is it possible to improve mathematical achievement by means of self-regulation strategies? Evaluation of an intervention in regular math classes. *European Journal of Psychology of Education, 24*(1), 17–31. doi:10.1007/BF03173472.

Rahimi, M., & Katal, M. (2012). Metacognitive strategies awareness and success in learning English as a foreign language: an overview. *Procedia - Social and Behavioral Sciences, 31*, 73–81.

Rizzo, P., Steinhausen, H. C., & Drechsler, R. (2010). Self-perception of self-regulatory skills in children with attention-deficit/hyperactivity disorder aged 8–10 years. *Attention Deficit and Hyperactivity Disorders, 2*(4), 171–83. doi:10.1007%2Fs12402-010-0043-x.

Rowe, M., Kairalla, J., & McCrae, C. (2009). Sleep in dementia caregivers and the effect of a nighttime monitoring system. *Journal of Nursing Scholarship, 42*(3), 338–347. doi:10.1111/j.1547-5069.2010.01337.x.

Schmitz, B. (2006). Advantages of studying processes in educational research. *Learning and Instruction, 16*, 433–449. doi:10.1016/j.learninstruc.2006.09.004.

Schmitz, P., & Perels, F. (2011). Self-monitoring of self-regulation during math homework behaviour using standardized diary tasks. *Metacognition and Learning, 6*, 255–273.

Schmitz, B., & Wiese, B. (2006). New perspectives for the evaluation of training sessions in self-regulated learning: Time-series analyses of diary task data. *Contemporary Educational Psychology, 31*, 64–96. doi:10.1016/j.cedpsych.2005.02.002.

Schmitz, B., Klug, J., & Schmidt, M. (2011). Assessing self-regulated learning using diary task measures with university students. In B. Zimmerman & D. Schunk (Eds.), *Handbook of self-regulation of learning and performance* (pp. 87–101). New York: Routledge.

Schunk, D. H., & Zimmerman, B. J. (Eds.). (1998). *Self-regulated learning: From teaching to self-reflective practice*. New York: Guilford Press.

Simard, D. (2004). Using diary tasks to promote meta linguistic reflection among elementary school students. *Language Awareness, 13*, 34–48. doi:10.1080/09658410408667084.

Singer, J. D., & Willett, J. B. (2003). Applied longitudinal data analysis: Methods for studying change and event occurrence. New York: Oxford University Press.

Stoeger, H., & Ziegler, A. (2008). Evaluation of a classroom based training to improve self-regulation in time management tasks during homework activities with fourth graders. *Metacognition and Learning, 3*, 207–230.

Stoeger, H., & Ziegler, A. (2010). Do pupils with different cognitive abilities benefit similarly from a self-regulated learning training program? *Gifted Education International, 26*, 110–123.

Stoeger, H., & Ziegler, A. (2011). Self-regulatory training through elementary-school students' homework completion. In B. Zimmerman & D. Schunk (Eds.), *Handbook of self-regulation of learning and performance* (pp. 87–101). New York: Routledge.

Velicer, W. F. (1976). Determining the number of components from the matrix of partial correlations. *Psychometrika, 41*, 321–327.

Whitebread, D., Coltman, P., Pino Pasternak, D., Sangster, C., Grau, V., Bingham, S., Almeqdad, Q., & Demetriou, D. (2009). The development of two observational tools for assessing metacognition and self-regulated learning in young children. *Metacognition and Learning, 4*(1), 63–85.

Whitty, M., Buchanan, T., Joinson, A., & Meredith, A. (2011). Not all lies are spontaneous: an examination of deception across different modes of communication. *Journal of the American Society for Information Science and Technology, 63*(1), 208–216. doi:10.1002/asi.21648.

Winne, P. (2011). A cognitive and metacognitive analysis of self-regulated learning. In B. Zimmerman & D. Schunk (Eds.), *Handbook of Self-Regulation of Learning and Performance* (pp. 87–101). New York: Routledge.

Wolters, C., Pintrich, P., & Karabenick, S. (March, 2003). *Assessing academic self-regulated learning*, paper presented at the Conference on Indicators of Positive Development: Definitions, Measures, and Prospective Validity. Sponsored by ChildTrends, National Institutes of Health.

Wong, F., Kember, D., Chung, L., & Yan, L. (1995). Assessing the level of student reflection from reflective journals. *Journal of Advanced Nursing, 22*, 48–57. doi:10.1046/j.1365-2648.1995.22010048.x.

Zimmerman, B. (1989). A social cognitive view of self-regulated academic learning. *Journal of Educational Psychology, 81*(3), 329–339. doi:10.1037//0022-0663.81.3.329.

Zimmerman, B. J. (2000). Attainment of self-regulation: A social cognitive perspective. In M. Boekaerts, P. Pintrich, & M. Zeidner (Eds.), *Handbook of self-regulation, research, and applications* (pp. 13–39). Orlando: Academic Press.

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

Zimmerman, B. (2013). From cognitive modeling to self-regulation: a social cognitive career path. *Educational Psychologist, 48*(3), 135–147.

Zimmerman, B., & Martinez-Pons, M. (1986). Development of a structured interview for assessing student use of self-regulated learning strategies. *American Educational Research Journal, 23*(4), 614–628. doi:10.3102/00028312023004614.