Fostering self-regulation in training complex cognitive tasks

Ludo W. van Meeuwen 1 · Saskia Brand-Gruwel 2 · Paul A. Kirschner 2 · Jeano J. P. R. de Bock 3 · Jeroen J. G. van Merriënboer 2, 4

Published online: 15 September 2017 © The Author(s) 2017. This article is an open access publication

Abstract In complex cognitive domains such as air traffic control, professionals must be able to adapt to and act upon continuing changes in a highly advanced technological work environment. To function optimally in such an environment, the controllers must be able to regulate their learning. Although these regulation skills should be part of their training, this is not usually the case. This study evaluates a training program that integrates air traffic control skills with regulation skills. The participants were 29 air traffic control students who followed either the original training program \( (n = 12) \) or a new program \( (n = 17) \) in which the development of regulation skills was embedded in the training of domain specific skills. Compared to students in the original program, the students in the new program showed increased self-efficacy in the use of self-regulated learning skills with improved performance in domain specific competences. The implications of these findings are discussed with regard to the daily training practice of complex cognitive skills.

Keywords Life-long learning · Complex cognitive task · Self-regulated learning · Self-efficacy · Continuous learning
Introduction

The nature of complex dynamic domains demands that experts in those domains keep up with developments in the relevant technologies and the regulations governing them. Air traffic control (ATC) falls into this domain. The continuous increase in air traffic (Eurocontrol Statfor 2010) is accompanied by changes in regulations for airline and air traffic safety, pollution, and noise abatement. To deal with these changes, technologies are developed and introduced into the running system with the requirement that human action will adapt perfectly. This combination requires professionals who, to maintain competency, are able to continue learning throughout their careers (Bolhuis 2003) despite the rapidly changing world around them (Jha et al. 2012; Van de Merwe et al. 2012; Van Merriënboer et al. 2009). Regulation skills increase the awareness of the experts regarding shortcomings in performance, which can motivate them to address these shortcomings in training programs (Eva and Regehr 2005, 2008). Therefore, in training a prospective air traffic controller, specific attention should be paid to the development of regulation skills that prepare them for continuous learning (Bolhuis 2003; Candy 1991). However, training does typically not focus on these skills, which raises two questions: How can students in a complex cognitive domain learn to use these skills while they learn complex domain-specific competencies? Will the use of these skills positively affect domain-specific learning outcomes? The objective of this study is to investigate a training design that integrates the development of both students’ regulation skills and their domain specific competences. The development of regulation skills and domain-specific competences are compared between an original training program and an integrated training program. Before we can answer the questions posed above, we need to elaborate on the complexity of ATC, define the regulation-of-learning skills that prepare for continuous learning and we should elaborate on training tools that can be designed to foster the development of these skills.

Task complexity

Solving an ATC task requires domain specific problem solving strategies which are mainly related to visual problem solving strategies (Van Meeuwen 2013; Van Meeuwen et al. 2014). As air traffic is dynamic, omission of decisions is not an option. This requires the development of many competences in information processing (i.e., perception, interpretation, dividing attention, planning and decision making) and action taking (i.e., communication, coordination, label and strip management, and equipment operation), which should lead to optimal ATC (i.e., safety and efficiency) regardless of outside influences (i.e., workload, teamwork, and others like motivation) (c.f., Oprins and Schuver 2003; Oprins et al. 2006).

Regulation of learning

Self-regulated learning (SRL) skills allow experts to critically evaluate their performance and choose learning activities that fit their learning needs in order to be prepared for future learning (Brand-Gruwel et al. 2014; Van Meeuwen et al 2013). SRL comprises strategies that can be classified according to Zimmerman’s five phases (1986, 1990): (1) orienting towards the task, (2) planning performance before starting the task, (3) monitoring task
performance while performing the task, (4) adjusting task performance, and (5) evaluating or self-assessing task performance after accomplishing the task. The evaluation of task performances should lead to setting new learning goals and selecting new learning materials to meet new set goals (Loyens et al. 2008). Knowles (1975) underscored the importance of these skills. He says that for directing own learning processes it is of importance that individuals take the initiative—with or without the help of others—in diagnosing their learning needs, formulating goals, identifying human and material resources, choosing and implementing appropriate learning strategies, and evaluating learning outcomes.

Because training in the use of SRL skills is a goal, it is important to be cognizant of the factors that foster the use and the development of these skills to improve students’ learning results (e.g., Winne and Azevedo 2014). Pintrich and De Groot (1990) showed that students’ self-efficacy is positively related to these self-regulation processes. Self-efficacy is one’s belief of being capable to carry out a task (Bandura 1997; Schunk 1985; cf., Zimmerman 2000). Self-efficacy is context specific. For this study the context that has been chosen is self-efficacy for learning a task, self-efficacy for performance on a task, and the appraisal of the task value in learning (Lodewyk and Winne 2005).

Self-efficacy and self-regulation can be considered as reciprocal causal factors. Several research (cf. Pajares 2008) supports the idea of Bandura’s model of triadic reciprocal causation (Bandura 1986) that self-efficacy and self-regulation bidirectional influence each other. As students’ self-efficacy and self-regulation processes have been found to be good predictors of performance (Pintrich and De Groot 1990) it is important when fostering students’ SRL skills to also take into account the development of students’ self-efficacy. However, doing so will have an influence on the design of a learning environment. We next describe how focusing on the development of these skills can be achieved by considering the training design.

### Training design to foster self-regulation

The design of a training program should meet the need to foster development of student self-efficacy and self-regulation in the training of complex cognitive skills. Based on Bandura’s (1986) model of triadic reciprocal causation, bidirectional interactions can be expected between the learning environment, students’ self-regulation activities, and students’ self-efficacy. Hence, if the design takes into consideration that mastery is a positive experience and students must learn to self-regulate their learning, the development of self-regulated learning activities can positively influence students’ self-efficacy and vice versa.

To improve learning outcomes, it is advisable to let students take initiative in the entire process of regulating their learning by giving them the opportunity to do so (Corbalan et al. 2006, 2008; Corbalan et al. 2010; Van Meeuwen et al. 2013). To prepare for future learning, students must be able to regulate and to evaluate their learning outcomes by carrying out self-assessments and diagnosing their learning needs in addition to learning how to formulate their learning goals and select resources for learning (i.e., learning tasks and coaching). However, such control over own learning processes requires SRL skills, which must be developed (Salden et al. 2006; Corbalan et al. 2010). Shared control over future learning can help students become more aware of their learning processes and of what is needed for future learning. In a shared control learning environment, the student and the system (e.g., coach) share the responsibility of future learning, and students
gradually learn to become aware of their learning needs and of matching them with possible task characteristics (Corbalan et al. 2010; Van Meeuwen et al. 2013). To self-regulate learning, students must learn how to optimally learn from a learning task and plan their actions accordingly. Monitoring and adjusting this performance helps them reach the defined learning goals. To give students positive experiences of mastery, learning tasks must be selected that optimally fit their individual learning needs. Shared control in which the responsibility of the control gradually shift from coach to student will have a positive effect not only on the SRL skills of the students, but will also benefit students’ self-efficacy because the guidance of the coach gives them more confidence during learning.

Development portfolio

A development portfolio can support SRL and stimulate continuous learning (Fig. 1) as well as facilitate shared control over selecting learning tasks that optimally fit individual learning needs (Kicken et al. 2008, 2009). The top of Fig. 1 shows the important elements of the development portfolio. The information about learning tasks as metadata (see next section) in combination with regulation prompts can give the student insight into their performance status. Based on the information in the portfolio, students are supported to self-regulate their learning. Hence, the development portfolio can foster the accurate selection of learning tasks that match the students’ learning needs. In this way the development portfolio is expected to support the increase of students’ self-efficacy (Zimmerman 2000).

Fig. 1 Elements in a development portfolio in relation to SRL, self-efficacy, and continuous learning
**Metadata**

For planning future learning, the portfolio comprises information about all available learning tasks (i.e., metadata), informing the user about the training characteristics of all tasks that can be used in planning future learning/performance. This metadata should contain details about all characteristics of the variety of tasks. The *four components instructional design model* (Van Merriënboer 1997) identified at least four characteristics that should be used in planning future learning. This information should thus be part of the metadata. First, a task is characterized by its complexity (i.e., number of elements and the interaction between the elements). In ATC, the complexity is mainly determined by the number of aircraft per unit of time and the number of potential conflicts, changes in runway use, and changes in weather conditions. In addition, non-nominal situations (e.g., emergencies) can increase task complexity. Second, the coach can provide support to lower a tasks’ training load. This support can vary from full support (i.e., worked examples) to no support (i.e., exam tasks). Third, specific supportive information can be required for the task (e.g., new call signs, specific communication by radio telephony, specific aircraft performances, etc.). If so, it can influence the task preparation process. Hence, the information should be studied in advance, which should thus be stated in the metadata. Fourth, the task can focus on training in a specific competence or skill (e.g., part task training), or it can be authentic and train for a range of competences. Therefore, the metadata should indicate which competences can be trained.

**Prompts**

To foster students’ self-regulation in learning, the development portfolio can contain learning task worksheets, including the metadata and *regulation prompts*. Research has shown that the use of regulation prompts is an effective way to train students’ SRL skills (e.g., Azevedo et al. 2016; Jossberger et al. 2010). At the right moments in the learning process, the development portfolio should prompt students to use the metadata to either orient themselves towards their future learning performance (prompts prior to the learning task) or evaluate their past learning performance (prompts after the learning task) and then plan the learning process.

**Hypotheses**

The main research question of this study is: What is the effect of a *training program* in which a development portfolio comprising metadata and regulation prompts is embedded and with shared control over the learning process on students’ SRL skills, self-efficacy, and task performance?

It is expected that this kind of training program would make it possible to involve students in regulating and delineating their learning (Corbalan et al. 2010; Van Meeuwen et al. 2013) and increase their regulation activities (Hypothesis 1). Specifically, it is expected that the training program would improve: student self-efficacy (self-efficacy in performance, self-efficacy in learning, task-value) (Hypothesis 1a) and student SRL skills (i.e., orientation to task, planning performance, monitoring performance, evaluating and self-assessment) (Hypothesis 1b). Moreover, because of better self-efficacy and SRL skills, we expect students in an integrated program to show better task performance than the students in a program focusing on ATC skills only (Hypothesis 2).
Method

Setting

This study was situated in the domain of ATC and focused on radar-based ATC training (i.e., area control surveillance; ACS). This comprises the training for area control of inbound and outbound air traffic at the Amsterdam Airport Schiphol and crossing aircraft to a flight level of 24,500 feet (approximately 7500 m). The ACS training program consists of 7 weeks of simulator training in which students in small groups of three or four perform tasks supervised by a coach. During this period, each student runs through 50 radar simulator tasks of 40 min, which provide the basic skills in area control surveillance. Training complexity increases in five main training steps. Each step comprises approximately 10 learning tasks (Fig. 2). For each learning task had a briefing and a debriefing.

Participants

The participants in this study were 29 students participating in an ACS course (Air Traffic Control the Netherlands). All participants had 9 months of training experience in ATC (age $M = 23.00$ years, $SD = 2.41$; 20 males, 9 females). For two and a half years, all regular students enrolled in the course on area control surveillance at Air Traffic Control the Netherlands participated in this study. During the first year, the original training program was run (original condition; $n = 12$; 10 males and 2 females). In the second and third years, the integrated training program was run (integrated condition; $n = 17$; 10 males and 7 females).

Materials

Original program

The original program comprised a 7 week training including 50 learning tasks per student. The learning tasks used a simulated airspace and would take approximately 35 min each. The coaches had a manual describing some metadata of each learning task (i.e., learning goals, traffic configurations, weather configurations, runway configurations, and coaching instructions). Prior to each learning task the coaches shared the corresponding metadata with the students and shared their ideas about optimal learning. After each task the coaches shared their feedback with the student.

Fig. 2 Intervention: an example of a training step
**Intervention**

The aim of the integrated training program was to foster students’ self-efficacy, and SRL, concurrently with the development of competence in ACS performance. In order to foster students’ self-efficacy, and SRL, different educational elements were embedded in the original training program (e.g., as described later and shown in Figs. 4, 5) based on Zimmerman (1990). The SRL elements have been shown to be a proven classification and also meet the preferences of the air traffic control trainers who indicated that they would be able to use this classification of skills in their education. The simulator learning tasks were redesigned and provided with more metadata (i.e., trainable competences, traffic load). Furthermore, process worksheets with all metadata were developed and provided to the students in order to increase their involvement. The worksheets gave the students learning task information (e.g., metadata) and prompted them for regulations before the learning tasks’ briefing and debriefing. The metadata and the worksheets yielded insight into the personal development of each student and comprised the development portfolio (Kicken et al. 2008, 2009). For an overview of the intervention, see Fig. 2.

**Worksheets and self-reports** A set of process worksheets and self-reports supported the intervention aimed to foster SRL, Self-efficacy and the development of ACS performance. This material prompted students both in preparation for the task and after completion of the task regarding regulations for the learning tasks’ briefing moments. Student SRL was prompted in three possible moments in the training: (1) prior to the learning task briefings (i.e., by learning task worksheets); (2) after each third learning task (i.e., by learning task self-reports), and (3) at the end of each training step, prior to the progression report meeting (i.e., by progression self-reports; PR-briefing). For an overview, see Fig. 2.

**Learning task worksheets** The learning task worksheets were divided into three parts (Fig. 3 shows a sample training-task worksheet). At the top of the sheet were metadata on training goals, traffic complexity (e.g., regional, inbound, and outbound crossing traffic) and task variables, such as weather conditions, runways in use, and training competencies (e.g., traffic flow, communication, perception). The second part provided preoperational regulation prompts, which asked students to think about individual training goals, how to reach them, and what they expected from the coach. Both parts prepared the student for the learning task. The third part included questions on performance and the implications for further training. The students answered these questions immediately after the learning task, which prepared them for the debriefing.

**Learning task self-reports** Students filled out the learning task self-reports every third learning task. The report was divided into two parts (Fig. 4). First, the report asked for an evaluation of the progress on 14 ATC main competences (described in the next subsection). The students were then asked to indicate the points of special interests that should be worked on in the next learning tasks, how to carry them out, and the support they expected from the coach. The self-report asked students to evaluate the progression of the prior three tasks by checking the individual learning task worksheets. This report also prepared the students for the learning task debriefing and therefore was filled out immediately after the learning task (i.e., instead of the last regulation prompt on the worksheet). Next, self-reports of two learning tasks provided data for one progress self-report.
Progress self-reports  Oprins and Schuver (2003) and Oprins et al. (2006) designed an ATC performance model to measure ATC performance. The model distinguishes factors in information processing (i.e., perception, interpretation, dividing attention, planning and decision making), actions (i.e., communication, coordination, label and strip management, and equipment operation), outcome (i.e., safety and efficiency), and influences (i.e., workload management, teamwork ability, and others like motivation). The application of this model allows the formulation of performance criteria on all four aspects and several sub-aspects. For example, safety is divided into three performance criteria: maintains separation minima correctly; builds in sufficient safety buffers; and switches from monitoring to vectoring in time (Oprins et al. 2006, p. 307). In this way, ATC performance of 14 competences can be scored based on 62 sub aspects of the performance criteria. The
progress self-reports assessed the 62 observable variables. It stimulated students to think about their progress on all assessment items. Students filled out the report at the end of each training step, which prepared them for the upcoming progress report meeting.

Coaching Coaching concentrated on briefing and debriefing the learning tasks and the progression report meetings. The learning task worksheets and self-reports ensured that the briefing and meetings focused on the learning challenges of the learning tasks. In the integrated program, the worksheets and self-reports prepared the students for the meetings.
Consequently, they were able to contribute from their point of view and experience control over their own learning process, instead of only receiving information from their coach.

**Comparison between programs**

In the original program, students were not equipped with the development portfolio (i.e., learning task worksheets, learning task self-reports, progression self-reports), so they went through their training program without a metadata overview, and were not prompted for regulation by the worksheets. Apart from the intervention itself and an update on airspace, all other aspects, such as training period (i.e., 7 weeks), number of learning tasks per student (i.e., 50), examination (according to the Eurocontrol Specification 2008), difficulty of pre- and post-test assignments (i.e., the number of aircraft and number of potential conflicts), and the moments of coaching were similar.

**Self-efficacy questionnaire**

A Dutch translation of the “Self and Task Perception Questionnaire” (STPQ; Lodewyk and Winne 2005) was used to measure self-efficacy. It was translated and validated with the permission of the authors. The original English version of this scale comprises 20 items with five answer options (e.g., fully disagree, disagree, disagree/agree, agree, fully agree). The self-efficacy items were based mainly on the motivated strategies for learning questionnaire (Pintrich et al. 1991). After translation into Dutch, the scale was re-translated into English to confirm the correct interpretation of the items. No differences in interpretation were found. Next, the Dutch translation was administered to 80 candidates in ATC training in The Netherlands (mean age = 23.4 years, SD = 2.40; 66 males, 14 females). The confirmatory factor analysis resulted in three constructs with 18 items in total: (1) measuring the sense of task agency, that is, self-efficacy for performance (6 items; Cronbach’s alpha = 0.83) (e.g., Knowing the difficulty of this project, the teacher, and my skills, I think I will do well on this task; I expect to do well on this task); (2) measuring the sense of future mastery, that is, ‘self-efficacy for learning’ (7 items; alpha = 0.73) (e.g., I’m confident I am learning the basic ideas in this task; I am enjoying the learning in this task), and (3) measuring the personal interest in the task, that is, ‘task value’ (5 items; alpha = 0.62) (e.g., Understanding the material of this task is important to me; I am interested in the material of this task).

**ATC assignment**

Two 10-min simulator tasks in ATC were designed for each program. The first task fits the expected level of performance at the start of the course (i.e., the number of aircraft was low with a minimum number of conflicts). The second task fits the expected performance level at the end of the course (i.e., a high number of aircraft with several possible conflicts ahead). These tasks contained both inbound and outbound traffic in the simulated environment of area control. The task was to maneuver the air traffic safely and efficiently to the indicated destinations. One experienced coach observed assignment performance. The tasks were divided into two phases: in the first minute, enough information was provided to allow students to orient to and plan the task. The performance phase was carried out in the remaining 9 min.
Cued retrospective report

Cued retrospective reporting (CRR) was used to measure the students’ SRL-levels (Van Gog et al. 2005). During the 10-min ATC assignment, student eye-movements were recorded with a Tobii 1750 eye-tracker that was connected to the ATC simulator. After the assignment, students watched their eye movements superimposed on the recording of the moving traffic situation. The eye movements and the traffic situation were played back at 75% of the actual speed. While they watched the replay, the students were asked to verbalize the thoughts they had during the assignment. If they were silent, they were prompted to keep talking about what they thought. The so called “gaze replay” showed the eye fixations based on the standard Tobii® studio fixation filter. A 25% delayed replay of the audio from the radio-telephony interaction supplemented the visual cue. The verbalizations from the CRR were recorded and transcribed. The CRR data from one student in the original condition were missing because of a recording error.

Coding scheme for SRL

An inductive-deductive method was used to develop the coding scheme to measure the amount of SRL reported in the CRR (an overview of the coding scheme is shown in Table 1). This means that first a coding scheme was designed based on literature. Next to this it was required to study the transcriptions and fine-tune the coding scheme. The scheme was based on the five components of Zimmerman’s SRL theory (1986, 1990): orientation, planning, monitoring, adjusting, and evaluation. Studying the transcriptions, however, showed the need to divide the evaluation category into five sub-categories: reflective evaluations (e.g., I have been thinking a while which flight level to give this aircraft); error analysis (i.e., negative evaluations; e.g., I directly felt punished that I had set about that conflict so clumsily); positive evaluations (e.g., I had that part under safe control, I could easily close my eyes for a while); learning evaluation (i.e., evaluative utterances on learning; e.g., I just tried something different. I put him as last in line, just to learn from the situation if something goes wrong). The use of CRR yielded retrospective evaluations on task performance (e.g., Retrospectively, I should have given him an expedite climb). Then, the coding scheme was used to score the frequency of students’ SRL utterances in their CRR. Therefore an interrater reliability analysis using the kappa statistic was performed on 10% of the transcriptions to determine consistency between two raters. Interrater reliability was acceptable (κ = 0.73); all remaining protocols were scored by one rater.

Scoring ATC assignment

To measure the ATC performance, 14 competences were scored based on the 62 sub aspects of the performance criteria, resulting in a final performance score between 0 and 100% (Oprins 2008).

Self-assessment accuracy

To measure the accuracy of self-assessment, the coach and the student assessed the ATC assignment performance on the 62 items assessment form for ATC competences (Oprins et al. 2006). Six of the 14 competences were most relevant for these short tasks (i.e., safety,
traffic flow, communication, mental model, planning, and decisiveness). To score the performance of the assignment, a 4-point Likert scale was used (1 = unsatisfactory; 2 = insufficient; 3 = sufficient; 4 = good). The measure of self-assessment accuracy was obtained by calculating the absolute difference between the coach’s score and the student’s score: the smaller the absolute difference (i.e., either overrated or underrated), the higher the self-assessment accuracy and vice versa.

Performance progress score

To measure the performance progress, two course assessments were available at the start and at the end of the ACS training period. The scores were based on the 62-item form for the assessment of all 14 ATC competences, which resulted in a performance progress score between 0 and 100% (Oprins et al. 2006).

Design and procedure

The measurements of the variables, the influence and development of self-efficacy, and SRL took place in individual pre- and post-60 min measurement sessions at the beginning and at the end of the training (see Fig. 5). The sessions were designed as follows: First, the students were told that the performance in this session would not influence their assessment in the ATC training. They were asked to answer some demographic questions. Next, they were informed about the start situation of the task, they received the corresponding flight strips (i.e., paper strips corresponding to the traffic containing relevant flight information), and they were allowed to orientate to the situation for 1 min. Based on the given information, they were asked to fill out the Dutch self-efficacy questionnaire in which the first

| Hypothesis | Main variable | Sub variable | Material |
|------------|---------------|--------------|----------|
| H1 (regulation) | Self-efficacy | For learning | Self-efficacy questionnaire |
| 1a | Self-efficacy | For task value | |
| 1b | Self-efficacy | For performance | |
| 1b | SRL | Orientation | Cued retrospective report and coding scheme |
| | | Planning | |
| | | Monitoring | |
| | | Adjusting | |
| | | Evaluation in action: reflection | |
| | | Evaluation in action: error analysis | |
| | | Evaluation in action: positive | |
| | | Evaluation on learning | |
| | | Evaluation on performance | |
| | | Self-assessment accuracy | 62-items assessment form |
| H2 (performance) | Performance progress score | | 62-items assessment from course assessment |

Table 1 Hypotheses and corresponding dependent variables and measurement materials
Fostering self-regulation in training complex cognitive tasks

A cued retrospective report concerned the first minute of orientation. The remaining 9 min of the task were then run, after which both the coach and the students assessed the performance separately. The session was closed with a CRR of the complete second part of the task.

The pre-measurement session took place at the end of the first simulator training step, followed by course assessment one. The post-measurement session took place in the fifth simulator training step, followed by course assessment two (For an overview, see Fig. 5).

Data analysis

To answer the question whether the intervention affected the students’ self-efficacy (Hypothesis 1a), SRL-skills (Hypothesis 1b), and performance (Hypothesis 2), the increases from the pretest to the posttest between students in the original condition and students in the integrated condition were compared. In all analyses, non-parametric independent sample tests were carried out. In the analyses reported here, unless indicated differently, a one-tailed significance level of 0.05 was used with \( N = 29 \) (i.e., original condition \( n = 12 \); integrated condition \( n = 17 \)) or less if indicated, due to missing values. The median rank (\( Mdn \)), the range of rank numbers, Mann–Whitney (\( U \)), \( z \)-score, level of significance (\( p \)), and effect size (\( r \)) are given.

Results

Self-efficacy

In order to test Hypothesis 1a (i.e., increase of self-efficacy), the differences in the increase of self-efficacy scores were compared between the original condition and the integrated condition. Table 2 provides the results. The analyses revealed no effects on increase of

![Fig. 5 Procedure](image)
self-efficacy for performance \((n = 29, U = 82.50, z = -0.871, p = 0.199, r = -0.16)\). However, a medium effect was found on increase of self-efficacy for learning. The increase in self-efficacy for learning in the integrated condition \((Mdn = 20, range 1–29)\) is significantly greater than in the original condition \((Mdn = 10, range 4–23, U = 57.50, z = -2.00, p = 0.023, r = -0.37)\). There was also an effect on increase of self-efficacy for task value. The increase in the integrated condition \((Mdn = 20, range 3–29)\) is significantly greater than in the original condition \((Mdn = 8.5, range 1–28, U = 37.50, z = -2.91, p = 0.001, r = -0.54)\), indicating that the intervention positively affected the increase of self-efficacy for learning and an increased interest in the task.

**Self-regulated learning**

In order to test Hypothesis 1b (i.e., increase of SRL including self-assessment), the increase of SRL scores between the original condition and the integrated condition was compared. The results summarized in Table 3 show a medium effect on increase of total SRL. The increase in total reported regulative utterances in the integrated condition \((Mdn = 19, range 5–28)\) was significantly greater compared to the original condition \((Mdn = 8, range 1–27, U = 39.50, z = -2.542, p = 0.006, r = -0.480)\). With respect to the increase of specific regulative activities, no effect between the two conditions was found on the increase of orientation \((U = 80.50, z = -0.619, p = 0.268, r = -0.117)\). A medium effect was found on the increase of planning activities: The integrated condition \((Mdn = 17, range 4–28)\) increased marginally significantly more than the original condition \((Mdn = 9, range 1–28, U = 60.00, z = -1.598, p = 0.055, r = -0.302)\). An effect was found on the increase of monitoring activities: The integrated condition \((Mdn = 19, range 6–28)\) increased significantly, compared to the original condition \((Mdn = 7, range 1–25, U = 34.50, z = -2.780, p = 0.003, r = -0.525)\). No effect between the two conditions was found on the increase of adjustment activities \((U = 67.50, z = -1.226, p = 0.110, r = -0.232)\). Effects were found on the increase of evaluation activities, including an effect on the increase of reflective evaluation activities in which the integrated condition \((Mdn = 20, range 8–28)\) increased significantly compared to the original condition \((Mdn = 6, range 1–18, U = 25.50, z = -3.216, p = 0.001, r = -0.610)\). A medium effect was found on the increase of error analysis activities in which the integrated condition \((Mdn = 18, range 10–28)\) increased significantly compared to the original condition.

Table 2 Means and standard deviations of increase of self-efficacy measures (STPQ)

|                      | Original condition \((n = 12)\) | Integrated condition \((n = 17)\) |
|----------------------|---------------------------------|----------------------------------|
|                      | \(M (SD)\)                      | \(M (SD)\)                       |
| Increase Pre Post    |                                 |                                 |
|                      |                                 |                                 |
| Self-efficacy for performance | 0.33 (2.19)                     | 1.24 (2.39)                      |
|                      | 19.67 (1.67)                     | 21.59 (1.87)                     |
|                      | 20.00 (2.37)                     | 22.82 (1.91)                     |
| Self-efficacy for learning** | -0.50 (1.09)                    | 0.65 (2.76)                      |
|                      | 27.92 (1.88)                     | 28.59 (1.70)                     |
|                      | 27.42 (1.24)                     | 29.24 (2.33)                     |
| Self-efficacy for task value** | -0.58 (1.98)                    | 1.18 (1.74)                      |
|                      | 19.33 (2.39)                     | 18.35 (1.36)                     |
|                      | 18.75 (1.14)                     | 19.53 (1.37)                     |

* \(p < 0.10\); ** \(p < 0.05\)
Table 3 Means and standard deviations of increase of self-regulated learning measures

|                                           | Original condition \((n = 11)\) | Integrated condition \((n = 17)\) |
|------------------------------------------|----------------------------------|-----------------------------------|
|                                           | \(M (SD)\)                       | \(M (SD)\)                       |
| Increase                                 | Pre                              | Post                             |
| Total regulative utterances**            | \(-14.73 (25.87)\)               | \(7.41 (16.66)\)                 |
| Orientation                              | \(-0.64 (2.16)\)                 | \(0.00 (2.50)\)                 |
| Planning*                                | \(-0.27 (2.05)\)                 | \(1.47 (2.81)\)                 |
| Monitoring**                             | \(-11.27 (12.36)\)               | \(1.18 (7.77)\)                 |
| Adjusting                                | \(-0.91 (9.10)\)                 | \(3.59 (6.87)\)                 |
| Evaluation in action: reflection**       | \(-0.73 (3.04)\)                 | \(4.00 (3.32)\)                 |
| Evaluation in action: error analysis**   | \(0.18 (1.25)\)                  | \(1.18 (1.29)\)                 |
| Evaluation in action: positive           | \(0.09 (0.83)\)                  | \(0.35 (0.86)\)                 |
| Evaluation on learning                   | \(-0.64 (1.29)\)                 | \(-0.29 (1.36)\)                |
| Evaluation on performance                | \(-0.81 (5.43)\)                 | \(-1.58 (4.76)\)                |

\* \(p < 0.10\); ** \(p < 0.05\)
condition \((Mdn = 6, \text{ range } 1–25, U = 48.50, z = -2.240, p = 0.013, r = -0.423)\). The results showed no effects between the two conditions on the increase of positive evaluation \((U = 76.50, z = -0.892, p = 0.187, r = -0.169)\), the increase of evaluation of learning \((U = 87.50, z = -0.291, p = 0.386, r = -0.055)\), and the increase of evaluation of performance \((U = 79.00, z = -0.684, p = 0.247, r = -0.129)\). These results indicate that the intervention program positively affected the increase of students’ SRL planning, monitoring, reflective evaluation, and error analysis skills but not the increase of their orientation, adjusting, positive evaluation, evaluation on learning, and evaluation on performance skills.

**Self-assessment accuracy**

Table 4 shows a summary of the results. The results show no differences between the conditions on the increase in self-assessment accuracy in safety \((U = 66.00, z = -0.859, p = 0.195, r = -0.17)\), traffic flow \((U = 87.00, z = -0.306, p = 0.380, r = -0.006)\), communication \((U = 80.00, z = -0.251, p = 0.401, r = -0.05)\), mental model \((U = 83.00, z = -0.494, p = 0.311, r = -0.09)\), planning \((U = 88.00, z = -0.259, p = 0.385, r = -0.05)\), and decisiveness \((U = 61.00, z = -0.876, p = 0.191, r = -0.17)\).

**Performance progress score**

In order to test Hypothesis 2 (i.e., increase of performance as measured by the course assessment), the differences in progress scores (i.e., differences between assessment one and assessment two) were compared between the original condition and the integrated condition. The results, which are summarized in Table 5, show a medium effect in the students’ performance progress scores. The integrated condition \((Mdn = 20, \text{ range } 2–29)\) showed significantly greater progress than the original condition did \((Mdn = 11, \text{ range } 1–26)\), \(U = 60.00, z = -1.863, p = 0.032, r = -0.35\).

**Discussion and conclusions**

The aim of this study was to investigate the implications of integrating the training of students’ regulation skills in a training program for domain specific skills in a complex cognitive domain. The results indicate that involving students in their learning process could result in improving regulation activities and better learning outcomes in complex cognitive tasks.

Hypothesis 1 stated that an integrated training program combining training in complex domain-specific competences and regulation skills increased both regulation activities and domain specific competences. The introduction of the integrated training program resulted in an increase in students’ self-efficacy (H1a), SRL (H1b) and an increase in students’ performance (H2). In line with Hypothesis 1a, an increase was shown in self-efficacy for learning (i.e., sense of future mastery) and self-efficacy for task value (i.e., personal interest in the task). Students in the integrated condition gained a more positive belief in how well they were prepared for the tasks, and they became more positive about their interest in the task. The sense of self-efficacy for task performance (i.e., task agency) did not differ between the two conditions. This means that the believe that one is capable of
|                              | Original condition |                     | Integrated condition |                     |
|------------------------------|--------------------|---------------------|----------------------|---------------------|
|                              | $M$ ($SD$)         | $n$                 | $M$ ($SD$)           | $n$                 |
|                              | Increase | lPrel | lPostl | n                     | Increase | lPrel | lPostl | n                     |
| Safety                       | 0.03 (0.52) | 0.48 (0.36) | 0.45 (0.42) | 11 | 0.09 (0.25) | 0.32 (0.25) | 0.43 (0.18) | 15 |
| Traffic flow                 | 0.03 (0.47) | 0.48 (0.39) | 0.51 (0.14) | 11 | 0.02 (0.37) | 0.38 (0.22) | 0.41 (0.17) | 17 |
| Communication                | -0.12 (0.33) | 0.49 (0.18) | 0.37 (0.32) | 10 | -0.13 (0.41) | 0.46 (0.34) | 0.33 (0.20) | 17 |
| Mental model                 | -0.18 (0.40) | 0.51 (0.38) | 0.33 (0.24) | 11 | -0.13 (0.44) | 0.45 (0.36) | 0.32 (0.25) | 17 |
| Planning                     | -0.11 (0.49) | 0.55 (0.43) | 0.44 (0.31) | 11 | -0.08 (0.35) | 0.44 (0.31) | 0.36 (0.26) | 17 |
| Decisiveness                 | 0.21 (0.59) | 0.61 (0.52) | 0.39 (0.25) | 11 | 0.02 (0.46) | 0.39 (0.28) | 0.41 (0.31) | 14 |
carrying out a task did not increase. Apparently, considering the definition of the first factor from the STPQ, the students did not experience sufficient task agency (Bandura 1982, 1997; Lodewyk and Winne 2005).

In line with Hypothesis 1b, greater development was found in the use of SRL skills in the integrated condition than in the original condition. The integrated condition showed more SRL skills than the original condition in terms of planning, monitoring, and self-evaluation (reflective evaluation and error analysis). This implies that the integrated condition increased in their ability to perceive own performance and recognize mistakes in that performance. Regarding the self-assessment accuracy, no differences were found between the conditions. This is an indication that improvement is needed in the quality of self-assessment and thus the instruction including feedback (cf. Butler and Winne 1995) should change to foster this development more effectively. For example according to the last mentioned authors’ model of feedback’s role in SRL.

In line with Hypothesis 2, the improvement shown in performance progress score is promising. In the integrated program, the development of self-efficacy and SRL skills was achieved simultaneously with an increase in learning outcomes. This result is in line with earlier research by Kicken et al. (2008, 2009) and Kicken et al. (2009), who studied the development of portfolio-based advice on task selection in a vocational training domain.

These conclusions acknowledge the training model (Fig. 2). The introduction of process worksheets has stimulated students to become involved in their learning process by letting them step away from the learning tasks and by making them use and improve their SRL skills.

Some limitations were caused by the application of the research to real practice. First, the relatively short training period of the ACS (i.e., 7 weeks) is a limitation, especially to prepare students for future learning. The worksheets involved students in their learning process, but they did not experience freedom in task selection. The delineation of learning needs and the translation to human and material resources for learning is of major importance for making choices for future learning (Knowles 1975). However, these skills can develop only when students experience choices in task selection and receive good feedback on this process (cf. Evans 2013). Third, the number of participants was limited. Fourth, because of circumstances, the ATC assignments, the learning tasks, and the course assessments used differed in two conditions because of updates in the simulated airspace used in training. Care was taken that the replacement ATC assignment and course assessments in the integrated condition were the same concerning complexity as those in the original condition. The results generally show comparable pre-test scores for both conditions. However, overall lower scores of the self-regulated learning measures can be found in the integrated condition. This is most probably due the updated simulated airspace

| Table 5 Means and standard deviations performance progress measure |
|----------------|----------------|----------------|
|                | Original condition (n = 12) | Integrated condition (n = 17) |
|                | M (SD) | M (SD) |
| Increase | Pre | Post | Increase | Pre | Post |
| Performance progress score** | 2.42 (10.82) | 87.67 (7.60) | 85.25 (8.82) | 4.53 (12.42) | 78.88 (9.47) | 83.41 (10.95) |

* p < 0.10; ** p < 0.05
but the eventual effects of the updated airspace will most probably have disappeared when the results of the pre- and post-measurements were subtracted. Fifth, the Cronbach Alphas for two of the three subscales of the SRL instrument are modest. A further development of SRL measurement scales can avoid this limitation in future studies. For these five reasons, it is necessary to be careful in generalizing the results (the performance progress scores in particular) to other situations.

Two aspects of the present study are important for future research. First, future research could focus on freedom in task selection in a shared control-task selection training design in combination with feedback (cf. Evans 2013). This is expected to foster not only the development of self-efficacy and SRL skills but could also foster the students’ ability to understand the consequences of learning needs for the selection of learning tasks. In future studies of task selection, students’ self-assessment skills will become a requirement for students. These skills would provide them with insight into their learning needs. Moreover, with better self-assessment skills, students will experience task mastery, which would foster their self-efficacy (Bandura 1982, 1997). Second, future research should include a longer training period. A virtuous circle can be expected when SRL and self-efficacy positively influence each other bi-directionally (Bandura 1986). A longer intervention might allow the further development of regulation skills, and allows monitoring the long-term influence of students’ regulation skills on the development of complex competences. This is in line with research showing larger impact of metacognitive effects in follow-up learning sessions than earlier measured short-term effects (Bannert et al. 2015).

To conclude, the results of this study imply that the design of an integrated learning program is appropriate to develop both domain specific competences and regulation skills as preparation for continuous learning (Bolhuis 2003; Candy 1991). The elements of a well-designed development portfolio played a crucial role in successful training. The study showed that providing students with relevant metadata and fostering them to prepare and evaluate their learning activities both improved the development of domain specific skills and fostered the development of self-regulation, which is a promising step towards improving the efficiency of training and continuous learning.

Acknowledgement This study was part of the future air-traffic management program funded by Knowledge Development Centre Mainport Schiphol. We would like to thank Mr. Wander Ziel for his participation in the organization of the measurements.

Compliance with ethical standards

Conflict of interest The authors hereby declare to comply with the ethical standards and author responsibilities as stated in the ETR&D journal instructions for authors. The authors declare that they have no conflict of interest.

Open Access This article is distributed under the terms of the Creative Commons Attribution 4.0 International License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made.

References

Azevedo, R., Martin, S. A., Taub, M., Mudrick, N. V., Millar, G. C., & Grafsgaard, J. (2016). Are pedagogical agents’ external regulation effective in fostering learning with intelligent tutoring system? In A. Micarelli, J. Stamper, & K. Panourgia (Eds.), ITS 2016 (p. 197). Switzerland: Springer International Publishing.

Bandura, A. (1982). Self-efficacy mechanism in human agency. American Psychologist, 37, 122–147.
Bandura, A. (1986). *Social foundations of thoughts and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice-Hall.

Bandura, A. (1997). *Self-efficacy: The exercise of control*. New York: W. H. Freeman.

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

Bolhuis, S. (2003). Towards process-oriented teaching for self-directed lifelong learning: A multidimensional perspective. *Learning and Instruction, 13*, 327–347.

Brand-Gruwel, S., Kester, L., Kicken, W., & Kirschner, P. A. (2014). Learning ability development in flexible learning environments. In J. M. Spector, D. M. Merrill, J. Elen, & M. J. Bishop (Eds.), *Handbook of research on educational communications and technology*. New York, NY: Springer.

Butler, D. L., & Winne, P. H. (1995). Feedback and self-regulated learning: A theoretical synthesis. *Review of Educational Research, 52*, 399–413.

Candy, P. (1991). *Self-direction for lifelong learning: A comprehensive guide to theory and practice*. San Francisco, CA: Jossey-Bass Publishers.

Corbalan, G., Kester, L., & Van Merriënoer, J. J. G. (2006). Towards a personalized task selection model with shared instructional control. *Instructional Sciences, 34*, 399–422.

Corbalan, G., Kester, L., & Van Merriënoer, J. J. G. (2008). Selecting learning tasks: Effects of adaptation and shared control on learning efficiency and task involvement. *Educational Psychologist, 33*, 733–756.

Corbalan, G., Van Merriënoer, J. J. G., & Kicken, W. (2010). Shared control over task selection: A way out of the self-directed learning paradox? *Technology, Instruction, Cognition and Learning, 8*, 119–136.

Eurocontrol Statfor. (2010). Long-term forecast flight movements 2010-2030. Retrieved from http://www.eurocontrol.int.

Eurocontrol Specification. (2008). Eurocontrol specification for the ATCO common core content initial training. Retrieved from http://www.eurocontrol.int.

Eva, K. W., & Regehr, G. (2005). Self-assessment in the health professions: A reformulation and research agenda. *Academic Medicine, 80*(10), S46–S54.

Eva, K. W., & Regehr, G. (2008). “I’ll never play professional football” and other fallacies of self-assessment. *Journal of Continuing Education in the Health Professions, 28*(1), 14–19.

Evans, C. (2013). Making sense of assessment feedback in higher education. *Review of Educational Research, 83*, 70–120.

Jha, P. D., Bisantz, A. M., Parasuraman, R., & Drury, C. G. (2012). Air traffic controllers’ performance in advance air traffic management system: Part I—performance results. *The International Journal of Aviation Psychology, 21*, 283–305.

Jossberger, H., Brand-Gruwel, S., Boshuizen, H. P. A., & Van de Wiel, M. (2010). The challenge of self-directed and self-regulated learning in vocational education: A theoretical analysis and synthesis of requirements. *Journal of Vocational Education & Training, 62*, 415–440.

Kicken, W., Brand-Gruwel, S., & Van Merriënoer, J. J. G. (2008). Scaffolding advice on task selection: A safe path toward self-directed learning in on-demand education. *Journal of Vocational Education and Training, 60*, 223–239.

Kicken, W., Brand-Gruwel, S., & Van Merriënoer, J. J. G. (2009a). Design and evaluation of a development portfolio: How to improve students’ self-directed learning skills. *Instructional Science, 37*, 453–473.

Kicken, W., Brand-Gruwel, S., Van Merriënoer, J. J. G., & Slot, W. (2009b). The effects of portfolio-based advice on the development of self-directed learning skills in secondary vocational education. *Educational Technology Research and Development, 57*, 439–460.

Knowles, M. S. (1975). *Self-directed learning*. Chicago: Follett Publishing Company.

Lodewyk, K. R., & Winne, P. H. (2005). Relations among the structure of learning tasks, achievement, and changes in self-efficacy in secondary students. *Journal of Educational Psychology, 97*, 3–12.

Loyens, S. M. M., Magda, J., & Rikers, R. M. J. P. (2008). Self-directed learning in problem based learning and its relationships with self-regulated learning. *Educational psychologist review, 20*, 411–427.

Oprins, E. (2008). Design of a competence-based assessment system for air traffic control training (Unpublished doctoral dissertation). Maastricht: Maastricht University.

Oprins, E., Burggraaff, E., & Van Weerdenburg, H. (2006). Design of a competence-based assessment system for air traffic control training. *The International Journal of Aviation Psychology, 16*, 297–320.

Oprins, E., & Schuver, M. (2003). Competentiegericht opleiden en beoordelen bij LVNL [Competence-based training and assessment at LVNL]. *HUFAG Nieuwsbrief, 6*, 2–4.
Pajares, F. (2008). Motivational role of self-efficacy beliefs in self-regulated learning: Theory, research and applications. In D. H. Schunk & B. J. Zimmerman (Eds.), *Motivation and self-regulated learning* (pp. 111–140). New York: Taylor & Francis Group.

Pintrich, P. R., & De Groot, E. V. (1990). Motivational and self-regulated learning components of classroom academic performance. *Journal of Educational Psychology, 82*, 33–40.

Pintrich, P. R., Smith, D. A., Garcia, T., & McKeachie, W. J. (1991). *A manual for the use of the Motivated Strategies for Learning Questionnaire (MSLQ)* (Technical Rep. No. 91-B-004). Ann Arbor, MI: University of Michigan, School of Education.

Salden, R. J. C. M., Paas, F., & Van Merriënboer, J. J. G. (2006). Personalised adaptive task selection in air traffic control: Effects on training efficiency and transfer. *Learning and Instruction, 16*, 350–362.

Schunk, D. H. (1985). Self-efficacy and school learning. *Psychology in the Schools, 22*, 208–223.

Van de Merwe, K., Oprins, E., Erikkson, F., & Van der Plaat, A. (2012). Model for task and job descriptions of air traffic controllers. *The International Journal of Aviation Psychology, 22*, 120–143.

Van Gog, T., Paas, F., Van Merriënboer, J. J. G., & Witte, P. (2005). Uncovering the problem-solving process: Cued retrospective reporting versus concurrent and retrospective reporting. *Journal of Experimental Psychology: Applied, 11*(1), 237–244.

Van Meeuwen, L. W. (2013). *Visual problem solving and self-regulation in training air traffic control (Unpublished doctoral dissertation)*. Heerlen: Open Universiteit.

Van Meeuwen, L. W., Brand-Gruwel, S., Kirschner, P. A., de Bock, J. J. P. R., Oprins, E., & Van Merriënboer, J. J. G. (2013). Self-directed learning in adaptive training systems: A plea for shared control. *Technology, Instruction, Cognition and Learning, 9*, 193–215.

Van Meeuwen, L. W., Jarodzka, H., Brand-Gruwel, S., Kirschner, P. A., de Bock, J. J. P. R., & van Merriënboer, J. J. G. (2014). Identification of effective visual problem solving strategies in a complex visual domain. *Learning and Instruction, 32*, 10–21.

Van Merriënboer, J. J. G. (1997). *Training complex cognitive skills*. Englewood Cliffs, NJ: Education Technology Publications.

Van Merriënboer, J. J. G., Kirschner, P. A., Paas, F., Sloep, P. B., & Cianiòls, M. C. J. (2009). Towards an integrated approach for research on lifelong learning. *Educational Technology Magazine: The Magazine for Managers of Change in Education, 49*, 3–15.

Winne, P. H., & Azevedo, R. (2014). Metacognition. In R. K. Sawyer (Ed.), *The Cambridge handbook of the learning sciences* (2nd ed.). Cambridge: Cambridge University Press.

Zimmerman, B. J. (1986). Development of self-regulated learning: Which are the key sub processes? *Contemporary Educational Psychology, 16*, 307–313.

Zimmerman, B. J. (1990). Self-regulated learning and academic achievement: An overview. *Educational Psychologist, 25*(1), 3–17.

Zimmerman, B. J. (2000). Self-efficacy: An essential move to learn. *Contemporary Educational Psychology, 25*, 82–91.

Ludo W. van Meeuwen is senior assessment consultant at the Educational and Students Affairs center of the Eindhoven University of Technology, the Netherlands.

Saskia Brand-Gruwel is full professor of Learning Sciences. Her chair is positioned within the Education & Training Institute of the Centre for Learning Sciences and Technologies, Open University, the Netherlands.

Paul A. Kirschner is Distinguished University Professor in educational research at the Open University, the Netherlands.

Jeano J. P. R. de Bock is program director at the Training, Design and Maintenance Department of Air Traffic Control, the Netherlands.

Jeroen J. G. van Merriënboer is full professor of Learning and Instruction and research program director at the Department of Educational Development and Research/School of Health Professions Education, Maastricht University.