Training self-assessment and task-selection skills: A cognitive approach to improving self-regulated learning

Citation for published version (APA):

Kostons, D., Van Gog, T., & Paas, F. (2012). Training self-assessment and task-selection skills: A cognitive approach to improving self-regulated learning. *Learning and Instruction*, *22*(2), 121-132. https://doi.org/10.1016/j.learninstruc.2011.08.004

DOI:
10.1016/j.learninstruc.2011.08.004

Document status and date:
Published: 01/04/2012

Document Version:
Peer reviewed version

Document license:
CC BY-NC-ND

Please check the document version of this publication:

• A submitted manuscript is the version of the article upon submission and before peer-review. There can be important differences between the submitted version and the official published version of record. People interested in the research are advised to contact the author for the final version of the publication, or visit the DOI to the publisher's website.
• The final author version and the galley proof are versions of the publication after peer review.
• The final published version features the final layout of the paper including the volume, issue and page numbers.

Link to publication

General rights
Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

• Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
• You may not further distribute the material or use it for any profit-making activity or commercial gain
• You may freely distribute the URL identifying the publication in the public portal.

If the publication is distributed under the terms of Article 25fa of the Dutch Copyright Act, indicated by the “Taverne” license above, please follow below link for the End User Agreement:

https://www.ou.nl/taverne-agreement

Take down policy
If you believe that this document breaches copyright please contact us at:

pure-support@ou.nl

providing details and we will investigate your claim.

Downloaded from https://research.ou.nl/ on date: 16 Feb. 2025

Open Universiteit
www.ou.nl

Learning and Instruction xxx (2011) 1–12

Contents lists available at SciVerse ScienceDirect

Learning and Instruction

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

Training self-assessment and task-selection skills: A cognitive approach to improving self-regulated learning

Danny Kostons[a,b,*], Tamara van Gog[c], Fred Paas[c,d]

[a] Centre for Learning Sciences and Technologies (CELSTEC), Open University, The Netherlands
[b] Groninger Institute for Educational Research, University of Groningen, The Netherlands
[c] Institute of Psychology, Erasmus University Rotterdam, The Netherlands
[d] Faculty of Education, University of Wollongong, Australia

ARTICLE INFO

Article history:
Received 19 January 2011
Received in revised form
25 August 2011
Accepted 26 August 2011

Keywords:
Self-regulated learning
Self-assessment
Task selection
Example-based learning

ABSTRACT

For self-regulated learning to be effective, students need to be able to accurately assess their own performance on a learning task and use this assessment for the selection of a new learning task. Evidence suggests, however, that students have difficulties with accurate self-assessment and task selection, which may explain the poor learning outcomes often found with self-regulated learning. In experiment 1, the hypothesis was investigated and confirmed that observing a human model engaging in self-assessment, task selection, or both could be effective for secondary education students' ($N = 80$) acquisition of self-assessment and task-selection skills. Experiment 2 investigated and confirmed the hypothesis that secondary education students' ($N = 90$) acquisition of self-assessment and task-selection skills, either through examples or through practice, would enhance the effectiveness of self-regulated learning. It can be concluded that self-assessment and task-selection skills indeed play an important role in self-regulated learning and that training these skills can significantly increase the amount of knowledge students can gain from self-regulated learning in which they choose their own learning tasks.

© 2011 Elsevier Ltd. All rights reserved.

1. Introduction

Self-regulated learning is an active, constructive process in which learners plan, monitor, and control their own learning process (e.g., Pintrich, 2000; Winne, 2001; Winne & Hadwin, 1998; Zimmerman, 1990; Zimmerman & Schunk, 2001). Self-regulated learning can occur at different levels, from learners controlling how long they engage in studying a given task or whether they want to restudy it (Karpicke, 2009; Metcalfe, 2009; Thiede & Dunlosky, 1999), to learners controlling what information they want to study (e.g., in a hypermedia learning environment; Azevedo, 2005; Azevedo & Cromley, 2004; Azevedo, Moos, Greene, Winters, & Cromley, 2008) or what learning tasks they want to work on (Corbalan, Kester, & Van Merriënboer, 2008; Kostons, Van Gog, & Paas, 2010; Ross, Morrison, & O'Dell, 1989). This article focuses on self-regulated learning in which learners can choose their own learning tasks. Research has shown that having control over what information to study or what tasks to work on is not effective for novices' self-regulated learning (Azevedo et al., 2008; Goforth, 1994; Lawless & Brown, 1997; Niemiec, Sikorski, & Walberg, 1996; Williams, 1996). We assume that this may be due to novices' lack of self-assessment and task-selection skills, which play a crucial role in this kind of self-regulated learning. To verify this assumption, we investigate whether training these skills results in higher self-assessment and task-selection accuracy (Experiment 1) and whether this in turn leads to higher learning outcomes attained through self-regulated learning (Experiment 2).

Providing learners with control over the learning tasks they work on is believed to foster their self-regulated learning skills and to result in personalized learning trajectories (Hannafin, 1984; Williams, 1996). Rather than having all learners follow the same instruction or practice schedule, which is often targeted at the average learner, personalized instruction allows learners who have difficulty with a task or topic to start at a lower level of complexity or obtain more support, while learners who find the new material easy can quickly move on to more complex materials. Such personalized instruction is expected to enhance students' motivation and learning outcomes more than non-personalized instruction that is the same for all students (Niemiec et al., 1996; Pintrich, 2004; Schnackenberg & Sullivan, 2000). However, there is little

* Corresponding author. Groninger Institute for Educational Research, University of Groningen, Grote Rozenstraat 3, 9712 TG Groningen, The Netherlands. Tel.: +31 503639609; fax: +31 503636670.
E-mail address: d.d.n.m.kostons@rug.nl (D. Kostons).

0959-4752/$ – see front matter © 2011 Elsevier Ltd. All rights reserved.
doi:10.1016/j.learninstruc.2011.08.004

evidence for both assumptions. First of all, research has shown that students do not apply and acquire self-regulation skills merely by engaging in self-regulated learning, but rather need additional training or instructional support such as prompts or tutoring (e.g., Aleven & Koedinger, 2002; Azevedo & Cromley, 2004; Van den Boom, Paas, & Van Merriënboer, 2007; Van den Boom, Paas, Van Merriënboer, & Van Gog, 2004). Secondly, although the assumption seems to be correct that personalized instruction can foster learning more compared to non-personalized instruction (e.g., Anderson, Corbett, Koedinger, & Pelletier, 1995; Camp, Paas, Rikers, & Van Merriënboer, 2001; Koedinger, Anderson, Hadley, & Mark, 1997; Salden, Paas, Broers, & Van Merriënboer, 2004), it is questionable whether self-regulated learning actually results in the adaptivity to students' needs that is required for effective personalized instruction.

When an instructional system is used to personalize instruction, it does so by monitoring and assessing a student's current level of knowledge and skill to select or suggest an appropriate next learning task. The assessment can be comprised of several aspects of students' performance (e.g., Anderson et al., 1995; Kalyuga & Sweller, 2004; Koedinger et al., 1997) or a combination of their performance and invested mental effort (e.g., Camp et al., 2001; Corbalan et al., 2008; Kalyuga, 2006; Salden et al., 2004). For self-regulated learning to be equally adaptive and effective, students should be able to accurately monitor and assess their own performance and recognize what an appropriate next task would be. However, there is quite some evidence that students, particularly novices who lack prior knowledge of the learning tasks, are not very accurate at monitoring, self-assessment, or task selection.

By monitoring one's own performance while working on a task, a student can construct a mental representation of the task performance process, which is considered to be a prerequisite for accurate self-assessment (e.g., Kostons, Van Gog, & Paas, 2009). However, both activities compete for limited working memory resources, which might become problematic under conditions of high cognitive load, as monitoring, task performance, or both may be negatively affected by a lack of resources (Van Gog, Kester, & Paas, 2011). Most learning tasks impose a high cognitive load, especially for novice learners (Sweller, Van Merriënboer, & Paas, 1998). When concurrent monitoring is hampered, learners will have a poor recollection of their performance (Kostons et al., 2009), which may hamper their self-assessment of that performance after the task.

But even with a good recollection of performance, accurate self-assessment is not guaranteed. Self-assessment may also be hampered by several biases that may cause learners to depend on the wrong kind of cues to assess their performance (for a review, see Bjork, 1999), such as hindsight bias (i.e., once an answer or solution procedure is known, e.g., after feedback, students are more likely to think that they could have produced it themselves), or availability bias (i.e., answers that come to mind easily are not only more likely to be provided but are also more likely to be assumed correct). Moreover, accurate self-assessment also seems to require some domain expertise (Dunning, Heath, & Suls, 2004; Dunning, Johnson, Erlinger, & Kruger, 2003). Individuals with higher levels of prior knowledge have been found to be more accurate self-assessors. This may be because their experience lowers the cognitive load imposed by the learning task (see Sweller et al., 1998), allowing them to devote more cognitive resources to monitoring their task performance, which likely provides them with a more accurate memory representation on which to base their assessment (Van Gog et al., 2011). In addition, learners may know that they should compare their performance to some standard, but do not know what criteria and standards their performance should meet (Dunning et al., 2004; Kostons et al., 2009). For example, a study by Tsivitanidou, Zacharia, and Hovardas (2011) investigated unsupported peer-assessment practices by students and the results showed that the overlap between assessment criteria formulated by the students themselves and assessment criteria determined by experts was very low. Individuals with more prior knowledge might be more accurate self-assessors because their experience provides them with more knowledge of the criteria and standards that good performance should meet (Dunning et al., 2003, 2004).

Inaccurate self-assessment, in turn, may negatively affect selection of an appropriate new learning task, for example, if students overestimate their performance, they may choose a task that is too difficult for them (cf. Azevedo, Guthrie, & Seibert, 2004; Moos & Azevedo, 2008a; Shapiro, 2004). Learners should ideally select tasks that are challenging, but not too difficult (cf. Vygotsky's, 1978, concept of 'zone of proximal development'). In practice, learners tend to select tasks that they are confident they can perform (i.e., self-efficacy, Bandura, 1997), but if those learners are over- or under-confident, selected tasks will not fit actual learning needs (Pajares & Kranzler, 1995; Stone, 1994). Moreover, even when self-assessment is accurate, novices may still experience problems in selecting appropriate learning tasks. When selecting a task, it is important to discern which aspects of a task are relevant for learning, such as the structural features of the task (e.g., type of task, complexity level, amount of support provided), and which aspects are less relevant, such as superficial cover stories (i.e., elements of the task story that are used to contextualize the problem, but do not alter the problem-solving procedure; see Fig. 1 for examples). Research has shown that novices may experience difficulties in discerning between these aspects (Chi, Glaser, & Rees, 1982; Quilici & Mayer, 2002; Ross, 1989) and tend to choose tasks based on irrelevant aspects (Ross & Morrison, 1989). When task selection is inaccurate, the chosen tasks are unlikely to fit the learner's level of prior knowledge or skills, so learners will end up working on tasks that are not aligned with their learning needs.

In sum, inaccuracies in self-assessment and task selection may lead to ineffective self-regulated learning when learners can select their own learning tasks. Support for this assumption comes from studies that have shown that providing novice learners with control over their learning process may have beneficial effects on their motivation or involvement (e.g., Corbalan et al., 2008; Schnackenberg & Sullivan, 2000), but has detrimental effects on learning outcomes when compared to teacher or computer controlled fixed or personalized instruction (see e.g., Azevedo et al., 2008; Lawless & Brown, 1997; Niemiec et al., 1996; Williams, 1996). Beneficial effects on learning outcomes attained through self-regulated learning have been found mainly for learners with higher levels of prior knowledge (Lawless & Brown, 1997; Moos & Azevedo, 2008b, 2008c; Niemiec et al., 1996; Scheiter & Gerjets, 2007; Schnackenberg & Sullivan, 2000; Steinberg, 1989), who, as mentioned above, are better able to monitor and assess their own performance than novices. In addition, Kostons et al. (2010) investigated whether secondary education students who differed in the amount of knowledge gained from studying in a self-regulated learning environment, also differed in self-assessment and task-selection skills. They found that students who gained more knowledge, were also more accurate self-assessors, and made better use of that assessment when selecting new tasks.

Given the important role that self-assessment and task-selection skills can be considered to play in the effectiveness of self-regulated learning, an important question is whether novice learners can be trained to become more accurate self-assessors and task selectors. This question is addressed in Experiment 1. If so, an even more important question, which is addressed in Experiment 2, is whether such training can improve the learning outcomes novices attain through self-regulated learning.

Please cite this article in press as: Kostons, D., et al., Training self-assessment and task-selection skills: A cognitive approach to improving self-regulated learning, *Learning and Instruction* (2011), doi:10.1016/j.learninstruc.2011.08.004

Complexity Level	Support Level	Learning Tasks				
Complexity 1 - 2 generations - 1 unknown - 1 solution - Deductive	Completion High support	Eye color	Hair structure	Shapes cat fur	Japanese Apple tree	Depression
	Completion Low support	Eye color	Hair structure	Sickle cell Anemia	Curve chicken beak	Guinea Pigs
	Conventional No support	Eye color	Hair structure	Huntington	Milk Allergy	Cleft Lip
Complexity 2 - 2 generations - 1 unknown - Multiple solutions - Deductive	Completion High support	Eye color	Hair structure	Flower color	Widow's peak	P.R.A.
	Completion Low support	Eye color	Hair structure	Shapes cat fur	Albinism	Pea plant
	Conventional No support	Eye color	Hair structure	Tongue Curling	Japanese Apple tree	Fruit flies
Complexity 3 - 2 generations - 1 unknown - Multiple solutions - Inductive	Completion High support	Eye color	Hair structure	Fruit flies	Curve Chicken beak	Wolfram syndrome
	Completion Low support	Eye color	Hair structure	Dog tail length	Japanese Apple tree	Milk allergy
	Conventional No support	Eye color	Hair structure	Freckles	Flower Color	Earlobes
Complexity 4 - 3 generations - 1 unknown - Multiple solutions - Both ways	Completion High support	Eye color	Hair structure	Albino	Shapes cat fur	Fruit flies
	Completion Low support	Eye color	Hair structure	Fruit flies	Tongue Curling	Flower color
	Conventional No support	Eye color	Hair structure	Pea plant	Dimples	Depression
Complexity 5 - 3 generations - 2 unknowns - Multiple solutions - Both ways	Completion High support	Eye color	Hair structure	Milk Allergy	Depression	Huntington disease
	Completion Low support	Eye color	Hair structure	Dog tail Length	Wolfram syndrome	Flower color
	Conventional No support	Eye color	Hair structure	Cystic Fibrosis	Fruit flies	Rat fur

Fig. 1. Task-selection database.

2. Experiment 1

Experiment 1 aimed to investigate whether an example-based training of self-assessment and task-selection skills led to higher accuracy of self-assessment and task selection. Research on *worked examples* inspired by cognitive theories such as ACT-R (Anderson, 1993; Van Merriënboer & Sweller, 2005) and cognitive load theory (Sweller, 1988; Sweller et al., 1998) and research on *modeling examples* inspired by social learning theory (Bandura, 1986) and cognitive apprenticeship (Collins, Brown, & Newman, 1989) has shown that both of these types of example-based learning are highly effective during the initial stages of problem-solving skill acquisition (for reviews, see Atkinson, Derry, Renkl, & Wortham, 2000; Van Gog & Rummel, 2010). Whereas worked examples are primarily based on a written account of a model's problem-solving procedure, modeling examples involve observing a model performing the task, which can take a variety of forms, not only live observation, but also watching a video in which the model is visible (e.g., Braaksma, Rijlaarsdam, & Van den Bergh, 2002), a video consisting of a screen capture of the model's computer screen in which the model is not visible (though s/he can be heard when a spoken explanation of what s/he is doing is provided; e.g., McLaren, Lim, & Koedinger, 2008; Van Gog, Jarodzka, Scheiter, Gerjets, & Paas, 2009), or an animation in which the model is represented by a pedagogical agent (e.g., Atkinson, 2002; Wouters, Paas, & Van Merriënboer, 2009).

Interesting from the perspective of the goal of our study, is that modeling examples have not only been used for teaching problem-solving skills (e.g., Schunk, 1981; Schunk & Hanson, 1985; McLaren et al., 2008) but also for improving *self-regulatory* skills, for example in trying to improve dart-throwing skills (Kitsantas, Zimmerman, & Cleary, 2000) or writing skills (Zimmerman & Kitsantas, 2002). In the study by Kitsantas et al., for example, students were shown either coping models (models who initially made many mistakes but gradually eliminated them) or mastery models (models who showed flawless performance from the start). While observing the coping models, students had to indicate whether the model made an error, and this was confirmed by the experimenter; if they did not notice an error, the experimenter informed them of it. Observation of a coping model led to higher dart-throwing skill, higher self-efficacy, higher intrinsic interest, and higher satisfaction with one's performance than observation of a mastery model, which led to higher skill, self-efficacy, interest, and satisfaction than having had no model. A similar pattern of results was found by Zimmerman and Kitsantas (2002) in their study of writing. These studies show that modeling examples can have positive effects on self-regulatory processes such as self-reactions (i.e., satisfaction), self-efficacy perceptions, and intrinsic interest.

However, the self-regulation measures these studies focused on were mostly affective in nature (i.e., self-reactions, self-efficacy, intrinsic interest). Furthermore, even though the coping models can be argued to model self-regulated learning, in the sense that

Please cite this article in press as: Kostons, D., et al., Training self-assessment and task-selection skills: A cognitive approach to improving self-regulated learning, *Learning and Instruction* (2011), doi:10.1016/j.learninstruc.2011.08.004

they started by making many errors but slowly reduced them, this self-regulated learning process was not addressed explicitly, and it was up to the students to notice the model's errors (i.e., there is no indication in the articles that the models verbalized awareness of errors or how to cope with them). So these studies did not investigate whether self-assessment or task-selection skills can be obtained through modeling of self-assessment or task selection. It could be argued that self-satisfaction could be a proxy for students' self-assessment (though it was assessed only once at the end of the entire test, instead of after each task), but that skill was not modeled. Moreover, task selection did not play a role in these studies.

Our study investigates whether explicit modeling of self-assessment and task-selection skills can improve these skills and foster self-regulated learning in which learners have control over the tasks they practice. We used modeling examples consisting of computer screen-recordings with spoken text to investigate our hypothesis that training secondary education students' self-assessment and task-selection skills would enhance the accuracy of those skills (hypothesis 1). By providing students with either no self-assessment and task-selection examples, only self-assessment examples, only task-selection examples, or both self-assessment and task-selection examples, we were not only able to investigate whether self-assessment and task selection can be trained, but also to determine whether an increase in accuracy in one skill would transfer to the other, for example, whether an increase in self-assessment accuracy would also lead to an increase in task-selection accuracy even if the latter was not modeled. Although we assume self-assessment outcomes to play an important role as input for task-selection, task selection constitutes a different skill and therefore we hypothesize that training one skill would not transfer, that is, would not lead to better performance on the other, at least when task selection performance is controlled for self-assessment accuracy (hypothesis 2).

2.1. Method

2.1.1. Participants and design

Participants were 80 Dutch students (44 female, 36 male; age $M = 15.23$, $SD = 0.53$) in their fourth year of pre-university education (i.e., the highest level of secondary education in the Netherlands with a six year duration) from a single school in the southern part of the Netherlands. All participants were used to working with computers; they had multiple lessons each week in which they utilized computers. A 2×2 factorial design was used with factors Self-Assessment Modeling Examples (Yes vs. No) and Task-Selection Modeling Examples (Yes vs. No). Participants were randomly assigned to one of the four conditions: (1) self-assessment and task-selection modeling examples (SA + TS; $n = 20$), (2) only self-assessment modeling examples (SA; $n = 20$), (3) only task-selection modeling examples (TS; $n = 19$), or (4) no self-assessment or task-selection examples (NO; $n = 21$). Students had not yet received any formal education on heredity, which is the subject of the learning materials used in this study.

2.1.2. Materials

2.1.2.1. Pretest and posttest. The pretest and posttest consisted of five paper and pencil heredity problems on the laws of Mendel, at five levels of complexity (for an example of a hereditary problem, see Appendix 1; for examples of test questions, see Appendix 2). The problems were presented in random order (i.e., not in order of complexity) and the order differed between the pretest and posttest. The posttest contained problems that were equivalent but not identical to the pretest problems: they had similar structural features but the surface features (cover stories) differed. The problems could be solved in five steps: (1) translating the phenotypes (i.e., expressions of genetic traits) described in the cover story into genotypes (a pair of upper and/or lower case letters representing genetic information); (2) putting these genotypes into a hereditary diagram; (3) determining the number of required Punnett Squares by looking at the direction of reasoning (deductive/inductive); (4) filling in the Punnett Square(s); and (5) extracting the final solution(s) from the Punnett Square(s). On both tests, participants were instructed to not only provide the final answer, but also write down the steps they took to reach the solution.

2.1.2.2. Mental effort rating. After each problem in the pretest and posttest, participants rated how much mental effort they had invested in solving the problem on a nine-point rating scale developed by Paas (1992) which ranges from (1) "very, very little effort" to (9) "very, very much effort". This single item is applied immediately after each learning task. It has been shown to be reliable and sensitive for differences in task complexity (i.e., the higher the complexity, the higher the mental effort ratings: Paas, Van Merrienboer, & Adam, 1994) and has been used frequently in the context of cognitive load research (for reviews, see Paas, Tuovinen, Tabbers, & Van Gerven, 2003; Van Gog & Paas, 2008). In this study, the mental effort ratings were additionally used for determining appropriate task selection in line with the method for system-controlled task selection used in the studies by Camp et al. (2001), Corbalan et al. (2008), Salden et al. (2004) and Salden, Paas, and Van Merriënboer (2006) where a system rather than the learner made task selections.

2.1.2.3. Self-assessment. After completing the mental effort scale following each task, participants self-assessed their performance on that task on a single, six point rating scale ranging from 0 to 5, assigning one point for each step in the problem-solving process they thought they had performed correctly, that is, (0) indicated "none of the steps correct" and (5) indicated "all steps correct". So self-assessment was not only operationalized as an evaluation of the outcome (step 5), but also as an evaluation of the problem-solving process, that is, of the steps taken to reach that outcome (steps 1–4). After the experiment, participants' performance on each task was scored by the experimenter on the same scale, also assigning one point for each solution step performed correctly, resulting in a maximum score of 25 on the pre- and posttest.

2.1.2.4. Task selection. After self-assessment, participants were informed what the complexity level was of the problem they had just worked on and were asked to indicate on an overview of the task database (see Fig. 1) what problem they would have selected next. Note though, that they did not get to work on that problem because the tasks in the pretest, modeling example phase, and posttest were the same for all students. The task database contained tasks at five levels of complexity (left column) and at each level, there were tasks that contained three levels of support: 1) high support: completion problems (i.e., partially worked-out examples that the learner has to complete, see Paas, 1992) with many steps already worked out and few for the learner to complete (white row); 2) low support: completion problems with few steps already worked out and many for the learner to complete (light gray row); and 3) no support: conventional problems which participants had to complete entirely (dark gray row). At each level of support within each complexity level, there were five tasks to choose from, which consisted of different cover stories (e.g., hair color, disorders; see Fig. 1). These cover stories did not influence the underlying structure of the problem-solving tasks, that is, all

Please cite this article in press as: Kostons, D., et al., Training self-assessment and task-selection skills: A cognitive approach to improving self-regulated learning, *Learning and Instruction* (2011), doi:10.1016/j.learninstruc.2011.08.004

problems within a complexity level could be solved using the same procedure.

2.1.2.5. Modeling examples. Participants were given four modeling examples consisting of digital videos of the model's computer screen recorded with Camtasia Studio, along with a spoken explanation by the model. The modeling examples were scripted by the authors and then recorded. The gender of the models varied: two examples were by two different male models, and the other two examples were by two different female models (because the model's gender might possibly influence students' learning by affecting self-efficacy; Schunk, 1989). The four examples either showed the model solving a heredity problem (NO), the model solving a heredity problem and assessing his or her own performance (SA), the model solving a heredity problem and selecting a new task based on a performance score that was presented as a given and not further explained (TS), or the model solving a heredity problem, assessing his or her own performance, and selecting a new task (SA + TS), depending on the assigned condition. The content of the examples was as follows:

(1) Problem solving (all conditions). The model performed the problem-solving task and provided a verbal explanation about which steps were taken, how to perform these steps, and why these steps had to be performed. Two models worked on problems of complexity level 1, and two models worked on problems of complexity level 2 (i.e., of the five complexity levels present in the task database and in the pretest and posttest; see Table 1). The quality of the models' performance varied between the examples: the first example showed a model accurately solving the problem, but in the other three examples the models made one or more errors (see Table 1). This was done to create variability in phases 2 and 3 of the examples, that is, in the model's self-assessment scores and task selections (i.e., if the model would not make any errors or would detect and correct them immediately, they would always have the highest possible self-assessment score). Following task performance, the models rated their invested mental effort on the nine point rating scale.
(2) Self-assessment (SA and SA + TS conditions): The models assessed their performance on the 6-point rating scale, providing a verbal explanation of how they were going about self-assessment and assigning themselves one point for each problem-solving step that had been performed correctly. The models' self-assessment was always accurate, that is, whereas some of the models made mistakes during problem-solving, none of them made any mistakes during self-assessment (i.e., if they had made an error during problem solving, they mentioned this during self-assessment and assigned themselves 0 points for that step). Participants were made aware of this through instruction on the screen at the start of the experiment.
(3) Task selection (TS and SA + TS conditions): The model selected a new task based on a combination of the performance score and the mental effort score, providing a verbal explanation of why both self-assessed performance and self-rated mental effort were taken into account in the selection, and how to go about selecting a new task using these two variables as input. The models used a table (see Fig. 2) in which the relationship between performance and mental effort scores was depicted. This table could be used to infer a recommended 'step size' for task selection. For example, a performance score of three, and a mental effort score of two, would result in a step size of +2, which essentially indicates the number of rows one is recommended to go back or progress to in the second column from the left in Fig. 1. A positive step size means a recommendation to select a more challenging task (i.e., less support or higher complexity level), a step size of zero means repeating a comparable task (i.e., same level of support and same complexity level), and a negative step size means a recommendation to select a simpler task (i.e., higher level of support or lower level of complexity). This kind of task-selection algorithm, based on performance and mental effort scores, has proven to lead to an effective learning path in studies on adaptive, personalized task selection (e.g., Camp et al., 2001; Corbalan et al., 2008; Kalyuga, 2006; Salden et al., 2004). The models' task selection was always accurate, that is, the models always chose a task in line with the rules for task selection we presented, and participants were made aware of this through instruction on the screen at the start of the experiment.

Table 1
Modeling example characteristics.

Example	Model	Problem-solving performance	Problem complexity level
1	Male 1	0 errors	Level 1
2	Female 1	2 errors	Level 1
3	Male 2	4 errors	Level 2
4	Female 2	1 error	Level 2

Participants in the NO condition observed the model's only performing the problem-solving task. In the time in which the participants in the other conditions observed the model's self-assessment and/or task selection, participants in the NO condition were instructed to indicate whether the model made any errors during task performance, and if so, what the errors were and what the correct step would have been. Finding and fixing errors in examples may foster the acquisition of problem-solving skills (see Große & Renkl, 2007), and it can also be expected to direct students' attention toward assessment of performance (of the model) to some extent.

2.1.3. Procedure

The experiment was conducted in a computer room at the participants' school in sessions of approximately 70 min, with 10–24 students per session. Prior to the experiment, participants had been randomly assigned to conditions. First, all participants completed the pretest on paper. They were given four minutes to complete each problem, followed by one minute for assessing their performance and selecting a next learning task (a previous study using the same problems had shown this to be sufficient time for solving conventional problems; Kostons et al., 2010). Students did not receive any feedback on whether their task-performance, self-assessment or task selections were correct. Participants were not allowed to proceed to the next problem before the time was up and time was kept by the experimenter using a stopwatch. After completing the pretest, participants studied the modeling

Performance			
4-5	+2	+1	0
2-3	+1	0	-1
0-1	0	-1	-2
	1, 2, 3	4, 5, 6	7, 8, 9
			Effort

Fig. 2. Determining task-selection step-size using performance and mental effort scores.

Please cite this article in press as: Kostons, D., et al., Training self-assessment and task-selection skills: A cognitive approach to improving self-regulated learning, *Learning and Instruction* (2011), doi:10.1016/j.learninstruc.2011.08.004

examples on the computer which varied according to their assigned condition (see materials section). Each participant had a headset for listening to the model's explanations. Finally, all participants completed the posttest on paper, according to a procedure similar to the pretest.

2.1.4. Data analysis

Self-assessment accuracy was determined by computing the absolute difference between participants' objective performance score and their self-assessed performance score. The lower this difference, the more accurate participants' self-assessment was (i.e., 0 = 100% accurate). For example, a student with a self-assessed score of 4, but an objective score of 1, would have a difference score of 3, which is fairly inaccurate, whereas a student with a self-assessed score of 4, but an objective score of 5, would have a difference score of 1, which is fairly accurate. Each participant's mean self-assessment accuracy was computed for the posttest, not including those problems on which both the objective and the subjective assessment were zero. This was done because we did not want to overestimate self-assessment accuracy; it is logical that students can state quite accurately that they were not able to solve a problem at all, and this is probably not very indicative of their self-assessment skill. This is also the reason why we will not analyze mean self-assessment and task-selection accuracy on the pretest: in both conditions there were too many problems that had not even been partially solved. We included self-assessment and the task-selection ratings in the pretest so that participants would get acquainted with them, as the models in the experimental conditions also used these and participants had to use them again during the posttest.

Task-selection accuracy was determined by first recoding subjective (i.e., self-assessed) performance scores and the mental effort scores indicated by the participants according to the table that was also used for the task-selection by the model (see Fig. 2), and then reading off the recommended step-size from this table. Recommended step size, essentially indicates the number of rows one is recommended to go back or progress to in the second column from the left in Fig. 1. For example, if a participant had completed the conventional problem at complexity level 2 and was recommended to progress two steps, this recommendation meant to choose a task with low support at complexity level 3. The absolute difference between the recommended step size and the step size chosen by the participant was then computed to indicate task-selection accuracy based on participants' self-assessed performance score. For five participants, task-selection accuracy could not be computed due to missing values in mental effort or task-selection data. One could also compute task-selection accuracy based on the objective performance scores, but we preferred to use the self-assessed performance score because this does not penalize participants for inaccuracies in their self-assessment, while using the objective performance score would.

2.2. Results

Data were analyzed using 2 × 2 ANOVAs with Self-Assessment Modeling Examples and Task-Selection Modeling Examples as factors, and the significance level set at 0.05. Partial eta squared (η_p^2) is reported as a measure of effect size, 0.01, 0.06, and 0.14 corresponding to small, medium, and large effect sizes, respectively. Note that exploration of our data indicated that not all variables were normally distributed. Therefore, we also performed non-parametric tests (Field, 2005; Siegel, 1957), which led to the same pattern of results with regards to the main effects as the parametric tests. Because both parametric and non-parametric analyses led to the same conclusions with regard to H_0 rejections, but we could not investigate interactions non-parametrically, we report the parametric data.

2.2.1. Pre-analysis

Pretest data from two participants were missing. Participants' mean performance on the pretest problems was 3.42 ($SD = 5.01$), and their mean performance on the posttest problems was 15.74 ($SD = 6.08$) indicating participants acquired procedural skills for solving heredity problems from the modeling examples. A 2 × 2 ANOVA with self-assessment modeling examples and task-selection modeling examples as independent variables and the pretest to posttest knowledge gain as dependent variable did not show any significant differences between conditions, $F(3, 74) < 1$, *ns*. There were also no significant correlations between individual's pre-test scores, pre-to-posttest learning gains, self-assessment accuracy or task-selection accuracy (all $p > 0.20$), indicating that these variables were independent from each other, as one would expect given this design.

2.2.2. Self-assessment accuracy

In line with hypothesis 1, a 2 × 2 ANOVA with self-assessment modeling examples and task-selection modeling examples as independent variables and self-assessment accuracy as dependent variable showed a significant main effect of the factor Self-Assessment Modeling Examples, indicating that participants who had studied self-assessment modeling examples were more accurate ($M = 0.81$, $SD = 0.54$) than participants who had not studied those examples ($M = 1.23$, $SD = 0.75$), $F(1, 76) = 8.04$, $MSE = 33.08$, $p = 0.006$, $\eta_p^2 = 0.10$. No main effect of Task-Selection Modeling Examples nor an interaction effect was found, both $F(1, 76) < 1$, *ns*, which is in line with hypothesis 2.

2.2.3. Task-selection accuracy

In line with hypothesis 1, a 2 × 2 ANOVA with self-assessment modeling examples and task-selection modeling examples as independent variables and task-selection accuracy as dependent variable task-selection accuracy showed a significant main effect of Task-Selection Modeling Examples, indicating that participants who had studied task-selection modeling examples were more accurate ($M = 2.31$, $SD = 2.18$) than participants who had not studied those examples ($M = 3.96$, $SD = 2.06$), $F(1, 71) = 11.57$, $MSE = 323.27$, $p = 0.001$, $\eta_p^2 = 0.14$. No main effect of Self-Assessment Modeling Examples, nor an interaction effect was found, both $F(1, 71) < 1$, *ns*, which is in line with hypothesis 2.

2.3. Discussion

Results of this experiment showed, in line with hypothesis 1, that students can acquire self-assessment and task-selection skills, which are considered to play a pivotal role in the effectiveness of self-directed learning, from studying modeling examples. In line with hypothesis 2, the results suggest that self-assessment and task-selection skills are different enough to justify the need for explicit training of both, as we found no indications (i.e., no interaction effects) that an increase in self-assessment accuracy also led to an increase in task-selection accuracy (or vice versa) when the latter was not modeled.

All students gained problem-solving skills from the examples, and there were no differences between conditions in problem-solving skill acquisition. For reasons of experimental control, all students received the same tasks on the posttest, so in this experiment they did not actually get to work on the tasks they had selected. A very important question, therefore, is whether students can apply the self-assessment and task-selection rules they acquired from studying modeling examples in a self-regulated

Please cite this article in press as: Kostons, D., et al., Training self-assessment and task-selection skills: A cognitive approach to improving self-regulated learning, *Learning and Instruction* (2011), doi:10.1016/j.learninstruc.2011.08.004

learning environment in which they are allowed to select which problems to work on. If they are able to do so, we would expect this to enhance the learning outcomes attained through self-regulated learning. This question is addressed in Experiment 2.

In this first experiment, we chose modeling examples as a means to train self-assessment and task-selection skills, because research has shown that example-based learning is a powerful instructional strategy. Thus far, in educational settings, examples have mostly been used for teaching cognitive skills, and this study adds further evidence that they are useful for teaching metacognitive skills as well (see also Kitsantas et al., 2000; Zimmerman & Kitsantas, 2002). A lot of research, especially on learning from worked examples, has demonstrated that for the acquisition of problem-solving skills instruction consisting of studying examples is more effective for novices than instruction consisting of practicing problem solving (see Atkinson et al., 2000; Sweller et al., 1998; Van Gog & Rummel, 2010). In this study, we did not investigate whether training self-assessment and task-selection skills via modeling examples was more effective than training those skills in some other way, for example via practice after having been explained the assessment and selection rules (i.e., how to come to a performance assessment score and how to combine performance and mental effort scores to select a new task). Therefore, Experiment 2 also investigated the effectiveness of examples compared to practice with self-assessment and task-selection rules.

3. Experiment 2

This experiment investigated the effects of teaching self-assessment and task-selection skills on the effectiveness of self-regulated learning. Learning outcomes attained after engaging in self-regulated learning will be compared for students who were not taught those skills, students who were taught those skills via modeling examples as in Experiment 1, and students who were taught those skills by explaining them the self-assessment and task-selection 'rules' and then allowing them to practice application of these rules by having them assess the model's performance and subsequently select a new task for the model based on that assessment. This third condition thus involves a kind of 'peer-assessment'. Peer-assessment is often implemented not only as a grading procedure, but also as a means to foster the development of both content knowledge and assessment skills (Dochy, Segers, & Sluijsmans, 1999). It has been suggested that engaging in peer-assessment activities might also develop self-assessment skills (Dochy et al., 1999; Somervell, 1993). However, only a few studies have tried to empirically demonstrate that assessing a peer's performance may subsequently improve self-assessment skills (see Oldfield & Macalpine, 1995; Searby & Ewers, 1997).

It is hypothesized that training self-assessment and task-selection skills (both via examples and practice) would result in higher self-assessment and task selection accuracy than obtained by the control condition (hypothesis 3a), but it is an open question as to whether examples or practice is more effective (question 3b). Moreover, it is hypothesized that such training leads to higher learning outcomes attained after self-regulated learning than obtained by the control condition (hypothesis 4a), but again, it is an open question as to which of the two training types (examples or practice) is more effective (question 4b). On the one hand, based on findings concerning the acquisition of problem-solving skills, we might expect example-based learning to be more effective for enhancing accuracy and learning outcomes. For example, in learning to use concept mapping as a learning strategy, it was found that examples (modeling) was more effective than practice, at least when students were asked to self-explain the examples (Hilbert & Renkl, 2009). On the other hand, in studies comparing learning from worked examples and learning from practicing with problem solving, students are generally not provided with any information concerning how to solve the problem. In this study, they are first explained the rules before they have to apply them to the model's performance.

3.1. Method

3.1.1. Participants and design

Ninety Dutch students (50 female, 40 male; age $M = 14.66$, $SD = 0.71$) in their third year of Higher General Secondary Education (the second highest level of secondary education in the Netherlands, with a 5 year duration) participated in this experiment. They were from a single school in the southern part of the Netherlands. All participants were used to working with computers; they had multiple lessons each week in which they utilized computers. Students in participating classes were randomly assigned to one of three conditions a week before the study; however, several students were absent during the experiment, resulting in unequal distribution over conditions: (1) self-assessment and task-selection skills taught via modeling examples (Modeling; $n = 32$); (2) self-assessment and task-selection skills taught via practice (Practice; $n = 25$), and no teaching of self-assessment and task-selection skills (Control; $n = 33$). Students had not yet received any formal education on heredity, the subject used in the experiment.

3.1.2. Materials

3.1.2.1. Pretest and posttest. The pretest and posttest were the same as in Experiment 1.

3.1.2.2. Mental effort rating. The same 9-point rating scale was used as in Experiment 1.

3.1.2.3. Self-assessment/peer-assessment. The same 6-point rating scale was used as in Experiment 1.

3.1.2.4. Task selection. The same procedure was used as in Experiment 1. During the self-regulated learning phase that was added in Experiment 2, students could click on the task they wanted to perform in the overview of the database (see Fig. 1) that was visible on their computer screen and then received that task to work on.

3.1.2.5. Training conditions. The Control (i.e., no training) condition was the same as in Experiment 1: participants observed the video of the model performing the problem-solving task and then had to find and fix errors in the model's performance. The Modeling condition was the same as the SA + TS condition in Experiment 1. In the Practice condition, participants were first explained the self-assessment and task-selection rules that were also used by the modeling examples, but without a concrete example. That is, concerning assessment, it was explained that each of the five steps in the problem-solving process could be right or wrong and contributed equally to the total assessment score. Concerning task selection, it was explained how mental effort and performance scores could be combined to infer a recommended step size, which could then be used to determine the next task by going back or progressing the recommended number of rows in the database overview. Then, they observed the video of the model performing the problem-solving task, assessed the model's performance, and selected a new task for the model. They were reminded of the rules immediately before each assessment and task selection moment during the training phase.

3.1.2.6. Self-regulated learning. In this experiment, a self-regulated learning phase was added after the training and prior to the

Please cite this article in press as: Kostons, D., et al., Training self-assessment and task-selection skills: A cognitive approach to improving self-regulated learning, *Learning and Instruction* (2011), doi:10.1016/j.learninstruc.2011.08.004

posttest, in which students could perform eight practice tasks of their own choice, in an electronic learning environment (see also Kostons et al., 2010) consisting of a Web application with a database connected to it that contained all learning tasks (see Fig. 1). The environment logged participants' answers to the problems, responses on mental effort and self-assessment rating scales, and task-selection choices.

In the self-regulated learning phase, participants went through the cycle of selecting a task, performing the task, rating their mental effort, and assessing their performance eight times. Based on Kostons et al. (2010) a maximum of five minutes was allotted per task. A visual and auditory warning was given when only one minute was left, and the system automatically continued to the self-assessment phase once time was up. One restriction on task selection was implemented of which participants were informed beforehand and were reminded of each time they went to the task-selection screen: They needed to complete at least one conventional task within a complexity level (correctly or incorrectly) before they could proceed to a higher complexity level (cf. Corbalan et al., 2008; Kostons et al., 2010). This rule had been implemented for two reasons. First, this should lead students to avoid choosing only tasks with high levels of support, but rather test themselves whether they could perform the task without any support, as they would have to do on the test. Second, this led to at least some conventional tasks being performed during the self-regulated learning phase, allowing for analyses on self-assessment and task-selection accuracy during self-regulated learning (see data analysis section). This rule had also been explained in the task-selection training (i.e., in the modeling examples and practice conditions).

3.1.3. Procedure

The experiment was run at the participants' school in sessions of approximately 110 min duration, with 9–20 students per session. Participants had been randomly assigned to conditions prior to the experiment. They first completed the pre-test, according to the same procedure used in Experiment 1. After the pretest, participants in the Modeling condition observed the four problem solving, self-assessment and task-selection modeling examples. Participants in the Practice condition received the explanation of the rules, observed the model solving the problem and then assessed the model's performance and selected a next task for the model four times. Participants in the Control condition observed the model solving the problem and then engaged in finding and fixing the mistakes in the model's demonstrated performance four times. Then, all participants engaged in self-regulated learning in the electronic learning environment, selecting learning tasks, performing those tasks, rating their mental effort, and self-assessing their performance eight times. Finally, participants proceeded to the posttest, completed according to the same procedure as in Experiment 1.

3.1.4. Data analysis

Self-assessment and task-selection accuracy on the posttest and during the self-regulated learning phase were determined according to the same procedure used in Experiment 1. Mean task-selection accuracy could not be computed for 31 participants due to missing values in either mental effort scores or task selection choices (these participants were rather equally distributed across conditions: Modeling: $n = 10$, Practice: $n = 9$, Control: $n = 12$). During the self-regulated learning phase, self-assessment and task-selection accuracy was only computed for conventional problems, because the completion problems provided support consisting of worked-out steps, which had consequences for self-assessment (i.e., assigning oneself a point for those steps, is not a judgment of whether one has correctly performed it, but of whether one is confident that one could correctly perform it) and because of the task-selection rule, restrictions in task selection applied as well. Note that not all participants performed conventional problems, and that the mean accuracy may be based on different numbers of conventional problems for different participants.

There were no significant correlations between individual's pretest scores, pre-to-posttest learning gains, self-assessment accuracy or task-selection accuracy (all $p > 0.20$), indicating that these variables were independent from each other.

3.2. Results

Means and standard deviations of pretest and posttest performance, learning gains, self-assessment, and task-selection accuracy per condition are provided in Table 2. ANOVAs with planned contrasts (significance level of 0.05) were used to test our hypotheses. Cohen's d is provided as a measure of effect size, with 0.25, 0.50, and 0.80 corresponding to small, medium, and large effect sizes, respectively. As in Experiment 1, some variables were not normally distributed and therefore non-parametric tests were also performed, which led to the same pattern of results of main effects as the parametric tests reported here. Because both parametric and non-parametric analyses led to the same conclusions with regards to H_0 rejections we report the parametric data.

3.2.1. Self-assessment accuracy

In line with hypothesis 3a, an ANOVA with planned contrasts, with condition (modeling, practice or control) as independent variable and learning gains as dependent variable, showed that self-assessment accuracy on the posttest was higher in the Modeling condition than in the Control condition, $t(87) = 1.81$, $p = 0.037$, one-tailed, $d = 0.41$. In contrast to hypothesis 4a, there was no significant difference between the Control condition and the Practice condition, $t(87) < 1$, *ns*. In addition, the Modeling and the Practice conditions did not differ from each other, $t(87) < 1$, *ns* (see question 3b). There were no significant differences between conditions in self-assessment accuracy on the conventional problems completed during the self-regulated learning phase, all $t(64) < 1$, *ns* (hypotheses 4a and question 4b).

3.2.2. Task-selection accuracy

In contrast to hypothesis 3a, an ANOVA with planned contrasts, with condition (modeling, practice or control) as independent variable and learning gains as dependent variable, showed that there were no significant differences between conditions in task-selection accuracy on the posttest (all $t(56) < 1$, *ns*). However, there were significant differences between conditions in task-selection accuracy

Table 2
Means (and SD) of pretest and posttest performance and accuracy data per condition.

	Modeling	Practice	Control
	M (SD)	M (SD)	M (SD)
Pretest score (max. = 25)	3.69 (3.73)	3.00 (2.92)	3.91 (3.36)
Posttest score (max. = 25)	15.91 (6.42)	14.84 (5.92)	12.06 (7.07)
Learning gain (max. = 25)	12.22 (6.03)	11.84 (5.44)	8.15 (6.66)
Self-assessment accuracy posttest (lower = better)	1.41 (0.78)	1.55 (0.48)	1.71 (0.68)
Self-assessment accuracy self-regulated learning (lower = better)	1.52 (1.23)	1.22 (1.15)	1.47 (0.89)
Task-selection accuracy posttest (lower = better)	4.20 (1.49)	4.87 (2.13)	4.28 (1.87)
Task-selection accuracy self-regulated learning (lower = better)	2.28 (1.60)	2.25 (1.60)	3.90 (2.67)

Please cite this article in press as: Kostons, D., et al., Training self-assessment and task-selection skills: A cognitive approach to improving self-regulated learning, *Learning and Instruction* (2011), doi:10.1016/j.learninstruc.2011.08.004

on the conventional problems completed during the self-regulated learning phase. In line with our expectation (see hypothesis 3a), accuracy in the Modeling condition was higher than in the Control condition ($t(63) = 2.74$, $p = 0.004$, one-tailed, $d = 0.74$) and accuracy in the Practice condition was higher than in the Control condition ($t(63) = 2.52$, $p = 0.007$, one-tailed, $d = 0.75$) during self-regulated learning. The Modeling and the Practice conditions did not differ from each other ($t(63) < 1$, *ns*) (question 3b).

3.2.3. Learning gains

In line with hypothesis 4a, an ANOVA with planned contrasts, with condition (modeling, practice or control) as independent variable and learning gains as dependent variable, showed that the Control condition gained less knowledge than both the Modeling condition ($t(87) = 2.68$, $p = 0.005$, one-tailed, $d = 0.64$) and the Practice condition ($t(87) = 2.27$, $p = 0.013$, one-tailed, $d = 0.61$). Regarding open question 4b, the Modeling and the Practice conditions did not differ from each other ($t(87) < 1$, *ns*).

3.3. Discussion

This experiment showed that in line with our hypothesis 3a, training students' self-assessment and task-selection skills enhanced the effectiveness of self-regulated learning in terms of learning gains. Training via modeling and practice seemed to be equally effective (question 3b). One would expect the increase in learning gains in these conditions compared to the control condition to be due to increased self-assessment and task-selection accuracy. The results partially support this assumption, but are not unequivocal (hypothesis 4a and question 4b). Regarding self-assessment accuracy, the modeling condition indeed outperformed the control condition on the posttest, but the practice condition did not. Moreover, no differences between conditions in self-assessment accuracy during self-regulated learning were found. Regarding task-selection accuracy, no significant differences were found on the posttest, but the modeling and practice conditions significantly outperformed the control condition during self-regulated learning.

Compared to Experiment 1, the effects of training on self-assessment and task-selection accuracy were much more difficult to establish in this second study. Note that the number of tasks (i.e., conventional problems) on which self-assessment and task selection accuracy could be measured during the self-regulated learning phase was rather low and differed per participant. On the posttest, all tasks were again the same for all participants as in Experiment 1, but each participant's experiences during the interspersed self-regulated learning phase were different and this may have modified the effects of the training on self-assessment and task-selection accuracy, making potential differences that existed after training more difficult to measure by the time the posttest took place.

4. General discussion

Our main aim was to investigate whether self-assessment and task-selection accuracy can be increased through training and whether such training enhances the effectiveness of self-regulated learning in which students have control over which learning tasks to engage in. Experiment 1 demonstrated that training consisting of observation of human models who engage in self-assessment and task selection improved students' self-assessment and task-selection skills. Experiment 2 showed that training self-assessment and task-selection skills, either through modeling as in Experiment 1 or by being explained the rules and then practicing those skills by assessing the model's performance and selecting a new task for the model, indeed enhanced the effectiveness of self-regulated learning.

Previous research on improving self-regulated learning in hypermedia environments has shown that training can improve students' application of self-regulation activities such as monitoring or planning during task performance and that this can increase their learning outcomes (e.g., Azevedo & Cromley, 2004). Our study extends that research by focusing on training self-assessment and task-selection skills for self-regulated learning situations in which learners have control over which learning tasks to engage in. Prior research on modeling examples also suggested that modeling can improve certain self-regulatory processes (Kitsantas et al., 2000; Zimmerman & Kitsantas, 2002). However, these studies did not investigate whether processes important for self-regulated learning can be taught via modeling, that is, in those studies, the models did not explicitly show students how to effectively engage in processes important for self-regulated learning. Our studies extend this research by showing that modeling self-assessment and task-selection skills can lead to the improvement of those skills, and that this affects the effectiveness of self-regulated learning.

The training we implemented was relatively simple and focused primarily on teaching students the kind of rules for self-assessment and task selection that are also implemented in e-learning systems to personalize instruction (e.g., Camp et al., 2001; Corbalan et al., 2008; Kalyuga, 2006; Salden et al., 2004). Even though the exact content of the rules might differ between tasks and domains, especially for self-assessment (i.e., the scale on which performance is assessed, or how performance is assessed), the underlying principles that the task-selection rules convey, such as "when selecting a task, do not only regard your performance, but also the amount of effort you invested" and "if your performance was high and your invested mental effort was low, you can select a more complex task" are likely to be effective across tasks and domains. An interesting question for future research therefore, is whether acquired self-assessment and task-selection skills can transfer to other tasks in the same domain or even to other domains.

Another interesting question is whether training of self-assessment skills positively affects monitoring during task performance. That is, if students know what is important for assessing their performance, their monitoring might become more focused. It will probably still require some cognitive capacity that cannot be devoted to the learning task, but when monitoring is more focused this may be less detrimental for learning. It might even foster learning, as students might be able to adjust their performance in response to evaluations of certain steps made during task performance.

Finally, given the reported relationship between prior knowledge and accuracy of self-assessment, an interesting question would be whether students with higher levels of prior knowledge than the novices who participated in our studies would still benefit from training self-assessment and task-selection skills, or whether such training would be unnecessary or even harmful for them (i.e., an 'expertise reversal effect' might occur; Kalyuga, 2007; Kalyuga, Ayres, Chandler, & Sweller, 2003).

A limitation of the studies presented here that was already mentioned and discussed above, is that in Experiment 2 we did not manage to fully replicate the effects on accuracy that we found in Experiment 1. Another limitation is that this study focused solely on cognitive factors. It should be noted that the standard deviations on learning gains were quite large in all conditions. So even though conditions with training produced higher learning gains on average than the control condition, differences in learning gains within conditions could potentially be explained by differences in students' motivation or goal orientation (e.g., De Bilde, Vansteenkiste, & Lens, 2011; Kolovelonis, Goudas, & Dermitzaki, 2011; Pintrich, 2000; Schunk, 1990). Combining measures of

Please cite this article in press as: Kostons, D., et al., Training self-assessment and task-selection skills: A cognitive approach to improving self-regulated learning, *Learning and Instruction* (2011), doi:10.1016/j.learninstruc.2011.08.004

cognitive and affective variables in future studies might shed light on this issue. Furthermore, students did not receive any feedback on whether their self-assessments or task-selections were correct (cf. Butler & Winne, 1995; Kicken, Brand-Gruwel, Van Merriënboer, & Slot, 2009). Such feedback might further improve learning from examples (e.g., Stark, Kopp, & Fischer, 2011). Finally, this study did not assess longer-term effects of training self-assessment and task-selection skills on the effectiveness of self-regulated learning. Future research should investigate whether the effects of training fade over time and if so, at what intervals the training should be repeated.

Despite these limitations, our studies resulted in an important finding for educational practice, showing that a relatively simple training intervention aimed at improvement of self-assessment and task-selection skills can significantly increase the amount of knowledge or skills students gain from self-regulated learning in which learners can choose their learning tasks from a large set of tasks.

Appendix 1. Example of task at complexity level 2

Description

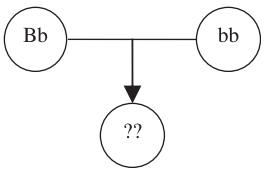
The eye color in humans is dependant on a gene that expresses itself dominantly as brown (B) and recessively as blue (b). Two parents produce an offspring. One of these parents has brown eyes and is heterozygote for this property. The other parent has blue eyes and is homozygote for this property. What genotypes could the offspring have?

Step 1: Translate to genotypes

Parent 1: Brown eyes. Brown is dominant so at least one capital B-allele. Heterozygote, meaning different alleles. Implies genotype of capital B and small b: Bb.

Parent 2: Blue eyes. Blue is recessive, so no capital B's allowed. Implies genotype of two small b's: bb.

Step 2: Put genotypes in hereditary tree



Step 3: Decide number of Punnett squares

Deductive reasoning, so 1 square.

Step 4: Create Punnett square(s)

	b	b
B	Bb	Bb
b	bb	bb

Step 5: Extract final answer

Answers: Bb and bb.

Appendix 2. Posttest (translated from Dutch)

Question 1 (complexity level 2)

Some people have a V-shape in their hair on top of their forehead, a so-called Widow's Peak. This shape is caused by a gene that expresses itself in its dominant form, and not in its recessive form. Two parents who both have a V-shape in their hair and are both heterozygote for this trait produce offspring. What are the possible genotypes for this offspring?

Question 2 (complexity level 3)

The shape of person's earlobe is determined by a gene, which in its dominant form expresses itself as an arch, and in its recessive form expresses itself as continuous. A parent with an arch, who is heterozygote for that trait, manages to produce a child that is homozygote for the trait and has continuous earlobes. What are the possible genotypes of the other parent?

Question 3 (complexity level 5)

Cystic Fibrosis (CF) is caused by a gene that in its recessive form expresses itself, but in its dominant form does not. Two parents produce offspring. One parent has CF, but the other parent's genotype is unknown. The offspring they produced also has CF. This offspring, together with a partner without CF who is homozygote for the trait, also produce offspring, whose genotype is unknown. What are the possible genotypes of the unknown parent and the unknown offspring?

Question 4 (complexity level 1)

A guinea pig's fur color is determined by a gene, which expresses itself as black in its dominant form and white in its recessive form. Two guinea pigs, who are both black and homozygote for that trait, produce offspring. What are the possible genotypes for this offspring?

Question 5 (complexity level 4)

The color of peas depends on a gene, which expresses itself as yellow in its dominant form and as green in its recessive form. Two peaplants, which are both yellow and heterozygote for that trait, produce offspring. No further information about that offspring is available. This offspring, with another peaplant that is green, manages to produce offspring that is yellow and heterozygote for that trait. What are the possible genotypes for the unknown peaplant?

References

Aleven, V. A. W. M. M., & Koedinger, K. R. (2002). An effective metacognitive strategy: learning by doing and explaining with a computer-based cognitive tutor. *Cognitive Science, 26*, 147–179. doi:10.1016/S0364-0213(02)00061-7.

Anderson, J. R. (1993). *Rules of the mind*. Hillsdale, NJ: Erlbaum.

Anderson, J. R., Corbett, A. T., Koedinger, K. R., & Pelletier, R. (1995). Cognitive tutors: lessons learned. *Journal of the Learning Sciences, 4*, 167–207. doi:10.1207/s15327809jls0402_2.

Atkinson, R. K. (2002). Optimizing learning from examples using animated pedagogical agents. *Journal of Educational Psychology, 94*, 416–427. doi:10.1037/0022-0663.94.2.416.

Atkinson, R. A., Derry, S. J., Renkl, A., & Wortham, D. (2000). Learning from examples: instructional principles from the worked examples research. *Review of Educational Research, 70*, 181–214. doi:10.3102/00346543070002181.

Azevedo, R. (2005). Using hypermedia as a metacognitive tool for enhancing student learning? The role of self-regulated learning. *Educational Psychologist, 40*, 199–209. doi:10.1207/s15326985ep4004_2.

Please cite this article in press as: Kostons, D., et al., Training self-assessment and task-selection skills: A cognitive approach to improving self-regulated learning, *Learning and Instruction* (2011), doi:10.1016/j.learninstruc.2011.08.004

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

Azevedo, R., Guthrie, G. T., & Seibert, D. (2004). The role of self-regulated learning in fostering students' conceptual understanding of complex systems with hypermedia. *Journal of Educational Computing Research, 30*, 87–111. doi:10.2190/DVWX-GM1T-6THQ-5WC7.

Azevedo, R., Moos, D., Greene, J. A., Winters, F. I., & Cromley, J. G. (2008). Why is externally-facilitated regulated learning more effective than self-regulated learning with hypermedia? *Educational Technology Research and Development, 56*, 45–72. doi:10.1007/s11423-007-9067-0.

Bandura, A. (1986). *Social foundations of thought and action: A social cognitive theory*. Englewood Cliffs, NJ: Prentice Hall.

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

Bjork, R. A. (1999). Assessing our own competence: heuristics and illusions. In D. Gopher, & A. Koriat (Eds.), *Attention and performance XVII. Cognitive regulation of performance: Interaction of theory and application* (pp. 435–459). Cambridge, MA: MITPress.

Braaksma, M. A. H., Rijlaarsdam, G., & Van den Bergh, H. (2002). Observational learning and the effects of model-observer similarity. *Journal of Educational Psychology, 94*, 405–415. doi:10.1037/0022-0663.94.2.405.

Butler, D. L., & Winne, P. H. (1995). Feedback and self-regulated learning: a theoretical synthesis. *Review of Educational Research, 65*, 245–281. doi:10.3102/00346543065003245.

Camp, G., Paas, F., Rikers, R., & Van Merriënboer, J. J. G. (2001). Dynamic problem selection in air traffic control training: a comparison between performance, mental effort and mental efficiency. *Computers in Human Behavior, 17*, 575–595. doi:10.1016/S0747-5632(01)00028-0.

Camtasia Studio (version 6) [Computer software]. Okemos, Michigan: Techsmith.

Chi, M. T. H., Glaser, R., & Rees, E. (1982). Expertise in problem solving. In R. J. Sternberg (Ed.), *Advances in the psychology of human intelligence, vol. 1*. Hillsdale, NJ: Erlbaum.

Collins, A., Brown, J. S., & Newman, S. E. (1989). Cognitive apprenticeship: teaching the crafts of reading, writing and mathematics. In L. B. Resnick (Ed.), *Knowing, learning, and instruction: Essays in honor of Robert Glaser* (pp. 453–494). Hillsdale, NJ: Erlbaum.

Corbalan, G., Kester, L., & Van Merriënboer, J. J. G. (2008). Selecting learning tasks: effects of adaptation and shared control on learning efficiency and task involvement. *Contemporary Educational Psychology, 33*, 733–756. doi:10.1016/j.cedpsych.2008.02.003.

De Bilde, J., Vansteenkiste, M., & Lens, W. (2011). Understanding the association between future time perspective and self-regulated learning through the lens of self-determination theory. *Learning and Instruction, 21*, 332–344. doi:10.1016/j.learninstruc.2010.03.002.

Dochy, F., Segers, M., & Sluijsmans, D. (1999). The use of self-, peer and co-assessment in higher education: a review. *Studies in Higher Education, 22*, 233–239. doi:10.1080/03075079912331379935.

Dunning, D., Heath, C., & Suls, J. M. (2004). Flawed self-assessment: implications for health, education, and the workplace. *Psychological Science in the Public Interest, 5*, 69–106. doi:10.1111/j.1529-1006.2004.00018.x.

Dunning, D., Johnson, K., Erlinger, J., & Kruger, J. (2003). Why people fail to recognize their own incompetence. *Current Directions in Psychological Science, 12*, 83–87.

Field, A. (2005). *Discovering statistics using SPSS* (2nd ed.). London: Sage.

Goforth, D. (1994). Learner control = decision making + information: a model and meta-analysis. *Journal of Educational Computing Research, 11*, 1–26. doi:10.2190/X799-7FDG-H0VQ-TH9C.

Große, C. S., & Renkl, A. (2007). Finding and fixing errors in worked examples: can this foster learning outcomes? *Learning and Instruction, 17*, 612–634. doi:10.1016/j.learninstruc.2007.09.008.

Hannafin, M. J. (1984). Guidelines for using locus of instructional control in the design of computer-assisted instruction. *Journal of Instructional Development, 7*(3), 6–10. doi:10.1007/BF02905753.

Hilbert, T. S., & Renkl, A. (2009). Learning how to use a computer-based concept-mapping tool: self-explaining examples helps. *Computers in Human Behavior, 25*, 267–274. doi:10.1016/j.chb.2008.12.006.

Kalyuga, S. (2006). Assessment of learners' organised knowledge structures in adaptive learning environments. *Applied Cognitive Psychology, 20*, 333–342. doi:10.1002/acp.1249.

Kalyuga, S. (2007). Expertise reversal effect and its implications for learner-tailored instruction. *Educational Psychology Review, 19*, 509–539. doi:10.1007/s10648-007-9054-3.

Kalyuga, S., Ayres, P., Chandler, P., & Sweller, J. (2003). Expertise reversal effect. *Educational Psychologist, 38*, 23–31. doi:10.1207/S15326985EP3801_4.

Kalyuga, S., & Sweller, J. (2004). Measuring knowledge to optimize cognitive load factors during instruction. *Journal of Educational Psychology, 96*, 558–568. doi:10.1037/0022-0663.96.3.558.

Karpicke, J. D. (2009). Metacognitive control and strategy selection: deciding to practice retrieval during learning. *Journal of Experimental Psychology: General, 138*, 469–486. doi:10.1037/a0017341.

Kicken, W., Brand-Gruwel, S., Van Merriënboer, J. J. G., & Slot, W. (2009). 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. doi:10.1007/s11423-009-9111-3.

Kitsantas, A., Zimmerman, B. J., & Cleary, T. (2000). The role of observation and emulation in the development of athletic self-regulation. *Journal of Educational Psychology, 92*, 811–817. doi:10.1037/0022-0663.92.4.811.

Koedinger, K. R., Anderson, J. R., Hadley, W. H., & Mark, M. A. (1997). Intelligent tutoring goes to school in the big city. *International Journal of Artificial Intelligence in Education, 8*, 30–43.

Kolovelonis, A., Goudas, M., & Dermitzaki, I. (2011). The effect of different goals and self-recording on self-regulation of learning a motor skill in a physical education setting. *Learning and Instruction, 21*, 355–364. doi:10.1016/j.learninstruc.2010.04.001.

Kostons, D., Van Gog, T., & Paas, F. (2009). How do I do? Investigating effects of expertise and performance-process records on self-assessment. *Applied Cognitive Psychology, 23*, 1256–1265. doi:10.1002/acp.1528.

Kostons, D., Van Gog, T., & Paas, F. (2010). Self-assessment and task selection in learner-controlled instruction: differences between effective and ineffective learners. *Computers & Education, 54*, 932–940. doi:10.1016/j.compedu.2009.09.025.

Lawless, K. A., & Brown, S. W. (1997). Multimedia learning environments: issues of learner control and navigation. *Instructional Science, 25*, 117–131. doi:10.1023/A:1002919531780.

McLaren, B. M., Lim, S., & Koedinger, K. R. (2008). When and how often should worked examples be given to students? New results and a summary of the currents state of research. In B. C. Love, K. McRae, & V. M. Sloutsky (Eds.), *Proceedings of the 30th annual conference of the cognitive science society* (pp. 2176–2181). Austin, TX: Cognitive Science Society.

Metcalfe, J. (2009). Metacognitive judgments and control of study. *Current Directions in Psychological Science, 18*, 159–163. doi:10.1111/j.1467-8721.2009.01628.x.

Moos, D. C., & Azevedo, R. (2008a). 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.

Moos, D. C., & Azevedo, R. (2008b). Self-regulated learning with hypermedia: the role of prior domain knowledge. *Contemporary Educational Psychology, 33*, 270–298. doi:10.1016/j.cedpsych.2007.03.001.

Moos, D. C., & Azevedo, R. (2008c). Monitoring, planning, and self-efficacy during learning with hypermedia: the impact of conceptual scaffolds. *Computers in Human Behavior, 24*, 1686–1706. doi:10.1016/j.chb.2007.07.001.

Niemiec, R. P., Sikorski, C., & Walberg, H. J. (1996). Learner-control effects: a review of reviews and a meta-analysis. *Journal of Educational Computing Research, 15*, 157–174. doi:10.2190/JV1U-EQ5P-X2PB-PDBA.

Oldfield, K. A., & Macalpine, J. M. K. (1995). Peer and self-assessment at tertiary level – an experiential report. *Assessment in Higher Education, 20*, 125–132. doi:10.1080/0260293950200113.

Paas, F. (1992). Training strategies for attaining transfer of problem-solving skill in statistics: a cognitive-load approach. *Journal of Educational Psychology, 84*, 429–434. doi:10.1037/0022-0663.84.4.429.

Paas, F., Tuovinen, J. E., Tabbers, H., & Van Gerven, P. W. M. (2003). Cognitive load measurement as a means to advance cognitive load theory. *Educational Psychologist, 38*, 63–71. doi:10.1207/S15326985EP3801_8.

Paas, F., Van Merrienboer, J. J. G., & Adam, J. J. (1994). Measurement of cognitive load in instructional research. *Perceptual and Motor Skills, 79*, 419–430. doi:10.2466/PMS.79.5.419-430.

Pajares, F., & Kranzler, J. (1995). Self-efficacy and general mental ability in mathematical problem-solving. *Contemporary Educational Psychology, 20*, 426–443. doi:10.1006/ceps.1995.1029.

Pintrich, P. R. (2000). Multiple goals, multiple pathways: the role of goal orientation in learning and achievement. *Journal of Educational Psychology, 92*, 544–555. doi:10.1037/0022-0663.92.3.544.

Pintrich, P. R. (2004). A conceptual framework for assessing motivation and self-regulated learning in college students. *Educational Psychology Review, 16*, 385–407. doi:10.1007/s10648-004-0006-x.

Quilici, J. L., & Mayer, R. E. (2002). Teaching students to recognize structural similarities between statistics word problems. *Applied Cognitive Psychology, 16*, 325–342. doi:10.1002/acp.796.

Ross, B. H. (1989). Remindings in learning and instruction. In S. Vosniadou, & A. Rotony (Eds.), *Similarity and analogical reasoning* (pp. 438–469). Cambridge, MA: Cambridge University Press.

Ross, S. M., & Morrison, G. R. (1989). In search of a happy medium in instructional technology research: issues concerning external validity, media replications, and learner control. *Educational Technology Research and Development, 37*(1), 19–33. doi:10.1007/BF02299043.

Ross, S. M., Morrison, G. R., & O'Dell, J. K. (1989). Uses and effects of learner control of context and instructional support in computer-based instruction. *Educational Technology Research and Development, 37*(4), 29–39. doi:10.1007/BF02307719.

Salden, R. J. C. M., Paas, F., Broers, N. J., & Van Merriënboer, J. J. G. (2004). Mental effort and performance as determinants for the dynamic selection of learning tasks in Air Traffic Control training. *Instructional Science, 32*, 153–172. doi:10.1023/B:TRUC.0000021814.03996.ff.

Salden, R. J. C. M., Paas, F., & Van Merriënboer, J. J. G. (2006). A comparison of approaches to learning task selection in the training of complex cognitive skills. *Computers in Human Behavior, 22*, 321–333. doi:10.1016/j.chb.2004.06.003.

Scheiter, K., & Gerjets, P. (2007). Learner control in hypermedia environments. *Educational Psychology Review, 19*, 285–307. doi:10.1007/s10648-007-9046-3.

Please cite this article in press as: Kostons, D., et al., Training self-assessment and task-selection skills: A cognitive approach to improving self-regulated learning, *Learning and Instruction* (2011), doi:10.1016/j.learninstruc.2011.08.004

ARTICLE IN PRESS

Schnackenberg, H. L., & Sullivan, H. J. (2000). Learner control over full and lean computer-based instruction under differing ability levels. *Educational Technology Research and Development, 48*(2), 19–35. doi:10.1007/BF02313399.

Schunk, D. H. (1981). Modeling and attributional effects on children's achievement: A self-efficacy analysis. *Journal of Educational Psychology*, 93–105.

Schunk, D. H. (1989). Self-efficacy and achievement behaviors. *Educational Psychology Review, 1*, 173–208. doi:10.1007/BF01320134.

Schunk, D. H. (1990). Goal setting and self-efficacy during self-regulated learning. *Educational Psychologist, 25*, 71–86. doi:10.1207/s15326985ep2501_6.

Schunk, D. H., & Hanson, A. R. (1985). Peer models: Influence on children's self-efficacy and achievement. *Journal of Educational Psychology*, 77, 313–322. doi:10.1037/0022-0663.77.3.313.

Searby, M., & Ewers, T. (1997). An evaluation of the use of peer assessment in higher education. *Assessment & Evaluation in Higher Education, 22*, 371–383. doi:10.1080/0260293970220402.

Shapiro, A. M. (2004). How including prior knowledge as a subject variable may change outcomes of learning research. *American Educational Research Journal, 41*, 159–189. doi:10.3102/00028312041001159.

Siegel, S. (1957). Nonparametric statistics. *The American Statistician, 11*(3), 13–19. http://www.jstor.org/stable/2685679.

Somervell, H. (1993). Issues in assessment, enterprise and higher education: the case for self-peer and collaborative assessment. *Assessment & Evaluation in Higher Education, 18*, 221–233. doi:10.1080/0260293930180306.

Stark, R., Kopp, V., & Fischer, M. R. (2011). Case-based learning with worked examples in complex domains: two experimental studies in undergraduate medical education. *Learning and Instruction, 21*, 22–33. doi:10.1016/j.learninstruc.2009.10.001.

Steinberg, E. R. (1989). Cognition and learner control: a literature review, 1977-88. *Journal of Computer-Based Instruction, 16*(4), 117–124.

Stone, D. (1994). Overconfidence in initial self-efficacy judgments: effects on decision processes and performance. *Organizational Behavior and Human Decision Processes, 59*, 452–474. doi:10.1006/obhd.1994.1069.

Sweller, J. (1988). Cognitive load during problem solving: effects on learning. *Cognitive Science, 12*, 257–285. doi:10.1016/0364-0213(88)90023-7.

Sweller, J., Van Merriënboer, J. J. G., & Paas, F. (1998). Cognitive architecture and instructional design. *Educational Psychology Review, 10*, 251–296. doi:10.1023/A:1022193728205.

Thiede, K. W., & Dunlosky, J. (1999). Toward a general model of self-regulated study: an analysis of selection of items for study and self-paced study time. *Journal of Experimental Psychology: Learning, Memory, and Cognition, 25*, 1024–1037. doi:10.1037/0278-7393.25.4.1024.

Tsivitanidou, O. E., Zacharia, Z. C., & Hovardas, T. (2011). Investigating secondary school students' unmediated peer assessment skills. *Learning and Instruction, 21*, 506–519. doi:10.1016/j.learninstruc.2010.08.002.

Van den Boom, G., Paas, F., & Van Merriënboer, J. J. G. (2007). Effects of elicited reflections combined with tutor or peer feedback on self-regulated learning and learning outcomes. *Learning and Instruction, 17*, 532–548. doi:10.1016/j.learninstruc.2007.09.003.

Van den Boom, G., Paas, F., Van Merriënboer, J. J. G., & Van Gog, T. (2004). Reflection prompts and tutor feedback in a web-based learning environment: effects on students' self-regulated learning competence. *Computers in Human Behavior, 20*, 551–567. doi:10.1016/j.chb.2003.10.001.

Van Gog, T., Jarodzka, H., Scheiter, K., Gerjets, P., & Paas, F. (2009). Attention guidance during example study via the model's eye movements. *Computers in Human Behavior, 25*, 785–791. doi:10.1016/j.chb.2009.02.007.

Van Gog, T., Kester, L., & Paas, F. (2011). Effects of concurrent monitoring on cognitive load and performance as a function of task complexity. *Applied Cognitive Psychology, 25*, 584–587. doi:10.1002/acp.1726.

Van Gog, T., & Paas, F. (2008). Instructional efficiency: revisiting the original construct in educational research. *Educational Psychologist, 43*, 16–26. doi:10.1080/00461520701756248.

Van Gog, T., & Rummel, N. (2010). Example-based learning. Integrating cognitive and social-cognitive research perspectives. *Educational Psychology Review, 22*, 155–174. doi:10.1007/s10648-010-9134-7.

Van Merriënboer, J. J. G., & Sweller, J. (2005). Cognitive load theory and complex learning: recent developments and future directions. *Educational Psychology Review, 17*, 147–177. doi:10.1007/s10648-005-3951-0.

Vygotsky, L. S. (1978). *Mind in society: The development of higher psychological processes*. Cambridge, MA: Harvard University Press.

Williams, M. (1996). Learner control and instructional technologies. In D. Jonassen (Ed.), *Handbook of research on educational communications and technology* (pp. 957–983). New York: Scholastic.

Winne, P. H. (2001). Self-regulated learning viewed from models of information processing. In B. J. Zimmerman, & D. H. Schunk (Eds.), *Self-regulated learning and academic achievement: Theoretical perspectives* (pp. 153–189). Mahwah, NJ: Lawrence Erlbauw.

Winne, P. H., & Hadwin, A. F. (1998). Studying as self-regulated learning. In D. J. Hacker, J. Dunlosky, & A. C. Graesser (Eds.), *Metacognition in education and practice. The educational psychology series* (pp. 277–304). Mahwah, NJ: Lawrence Erlbaum.

Wouters, P., Paas, F., & Van Merrienboer, J. J. G. (2009). Observational learning from animated models: effects of modality and reflection on transfer. *Contemporary Educational Psychology, 34*, 1–8. doi:10.1016/j.cedpsych.2008.03.001.

Zimmerman, B. J. (1990). Self-regulated learning and academic achievement: an overview. *Educational Psychologist, 25*, 3–17. doi:10.1207/s15326985ep2501_2.

Zimmerman, B. J., & Kitsantas, A. (2002). Acquiring writing revision and self-regulatory skill through observation and emulation. *Journal of Educational Psychology, 94*, 660–668. doi:10.1037/0022-0663.94.4.660.

Zimmerman, B. J., & Schunk, D. H. (Eds.). (2001). *Self-regulated learning and academic achievement: Theoretical perspectives* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum.

Please cite this article in press as: Kostons, D., et al., Training self-assessment and task-selection skills: A cognitive approach to improving self-regulated learning, *Learning and Instruction* (2011), doi:10.1016/j.learninstruc.2011.08.004